



Norwegian University of
Science and Technology

User-centered methods for small game development studios

Amund Lieng

01-06-2016

Master's Thesis

Master of Science in Interaction Design

30 ECTS

Department of Computer Science and Media Technology
Norwegian University of Science and Technology,

Supervisor: Mariusz Nowostawski

Preface

This is a Master's thesis in MIXD (Master of Science in Interaction Design) at NTNU. It was written during the spring semester of 2016. The goal of the thesis is to identify methods that can be used by small game developers to test their games.

Parts of the thesis has been done in cooperation with *Suttung Digital* and the European Horizon 2020 CityCop project. These groups provided the author with testable prototypes. Tests were run on these prototypes and the results were shared.

This thesis is written for readers with a basic knowledge of user-centered design and usability.

01-06-2016

Amund Lieng

Acknowledgment

I would like to thank the following persons and companies for their help during the writing of this thesis:

Suttung Digital, for providing a working prototype of their game *Fancy Diving* and for their participation in a case study where they were interviewed.

Konstantinos Boletsis, for providing me with prototypes for two games for the *CityCop* project.

Mariusz Nowostawski, for supervising me during the writing of this thesis and for putting me in contact with Suttung Digital and the *CityCop* project.

A.L.

Abstract

This thesis takes a look at how to adapt user-centered methods for user testing with small game developer studios. Reviewing earlier work reveals that little effort has been put into developing user-centered methods for these developers.

Three methods based on earlier work have been used for testing games. This is done to analyze how well the methods perform. These methods have also met specific criteria. These criteria state that the methods must be usable with little preparation. They must be fast and easy to conduct. They can be used with small sample sizes. The selected methods are a heuristic evaluation, playtesting (a type of observation study), and interviews.

Four games are tested with the selected methods in the study. The results from the tests show that the methods are usable for testing the selected games. It is shown that there are differences in their effectiveness depending on the complexity of the game. Two of the games produce large amounts of usable data. These two games are more complex than the other two which are more conceptual in nature. These two games still produced valuable information that helps to guide the development.

A case study where Suttung Digital is interviewed some weeks after their game has been tested reveals that there is a need for these methods. The issues found during the testing of their game are substantial enough for them to need a few months to implement. The case study shows that the methods are understandable to an audience with little knowledge of usability testing and that a framework for how to conduct these tests is wanted.

Contents

Preface	i
Acknowledgment	ii
Abstract	iii
Contents	iv
List of Figures	viii
List of Tables	ix
1 Introduction	1
1.1 Problem description	1
1.2 Document structure	2
2 Literature review	3
2.1 What is User-centered design?	3
2.2 Why test games?	3
2.3 Why are games and game research important?	4
2.4 Does game testing differ from testing traditional software?	4
2.5 What can be tested in games?	5
2.6 Can existing heuristics be used?	6
2.7 Which methods have been developed for game testing?	9
2.8 Who should be recruited to participate in testing?	11
2.9 How many test participants are needed?	11
2.10 When should testing be conducted?	12
2.11 When should testing stop?	12
2.12 What has been done to meet the need of small game developers?	12
2.13 Which methods should be used by small game developers?	13
2.14 Can the use of UCD-methods affect the game design intent?	13
2.15 Usability vs. playability	14
2.16 How to rate issues	14
2.17 Summary	15
3 Methodology	16
3.1 Selection of methods	16
3.2 Evaluating the methods	17
3.3 Selection of participants	17
3.4 Results	18
3.5 Analysis and evaluation	18
4 Results and analysis	19
4.1 First Test - "Alex and the Search for the Elemental Stone"	19
4.1.1 Test description	19

4.1.2	The Game: "Alex and the Search for the Elemental Stone"	20
4.1.3	Comparison of results	23
4.1.4	Evaluation of methodologies	24
4.1.5	Other Notes	25
4.2	Second Test - "Fancy Diving"	26
4.2.1	Test description	26
4.2.2	The Game: "Fancy Diving"	27
4.2.3	Comparison of results	27
4.3	Evaluation of methodologies	28
4.4	Third Test - "CityCop Facesketch"	29
4.4.1	Test Description	29
4.4.2	The Game: "CityCop Facesketch"	29
4.4.3	Comparison of results	29
4.4.4	Evaluation of methodologies	30
4.5	Fourth Test - "CityCop Identiface"	31
4.5.1	Test Description	31
4.5.2	The Game: "CityCop Identiface"	32
4.5.3	Results	33
4.5.4	Evaluation of methodologies	33
4.6	Analysis of test results	34
5	Case Study - Interview with Suttung Digital	35
5.1	The interview	35
5.2	Summary	40
5.3	Outcome	41
6	Discussion	42
6.1	Problem statement	42
6.1.1	Ease of use	42
6.1.2	Competence requirement	42
6.1.3	Small sample sizes	43
6.2	Differences between the methods	43
6.3	Results	43
6.4	Maintaining artistic integrity	43
6.5	Limitations of the study	44
6.5.1	Demographic	44
6.5.2	Development cycle	44
6.5.3	Sample sizes and game complexity	44
6.5.4	Full user experience	45
6.5.5	Platform differences	45
6.5.6	Analyzing the results	45
6.5.7	Severity rating	45
6.5.8	Occurrence rate	46
6.6	Summary	46

7 Conclusion	47
7.1 Suggestion for further research	47
Bibliography	48
Glossary	51
A Alex and the Search for the Elemental Stone - Test Results	I
A.1 Heuristic Evaluation	I
A.1.1 Heuristics	I
A.1.2 Evaluation	III
A.1.3 Recommendations	V
A.2 Play session	VI
A.2.1 Test Participant 1	VI
A.2.2 Test Participant 2	VIII
A.2.3 Test Participant 3	X
A.2.4 Recommendations	XI
A.3 Post-interview questions based on heuristics	XIV
A.3.1 Questions and answers	XIV
A.3.2 Recommendations	XVII
B Fancy Diving - Test Results	XIX
B.1 Heuristic Evaluation	XIX
B.1.1 Heuristics	XIX
B.1.2 Evaluation	XX
B.1.3 Recommendations	XXII
B.2 Playtesting	XXIII
B.2.1 Test Participant 1	XXIII
B.2.2 Test Participant 2	XXIII
B.2.3 Test Participant 3	XXIV
B.2.4 Test Participant 4	XXIV
B.2.5 Test Participant 5	XXV
B.2.6 Recommendations	XXV
B.3 Post-interview	XXVII
B.4 Post-interview	XXVII
B.4.1 Questions and answers	XXVII
B.4.2 Recommendations	XXXII
C Fancy Diving - Trick phase redesign	XXXIV
C.1 Goals	XXXIV
C.2 QTE(Quick Time Event) mode	XXXIV
D CityCop - Facesketch - Test Results	XXXV
D.1 Heuristic Evaluation	XXXV
D.1.1 Heuristics	XXXV
D.1.2 Evaluation	XXXVI
D.1.3 Recommendations	XXXVII
D.2 Playtesting	XXXVIII

D.2.1	Test Participant 1	XXXVIII
D.2.2	Test Participant 2	XXXVIII
D.2.3	Test Participant 3	XXXVIII
D.2.4	Test Participant 4	XXXIX
D.2.5	Test Participant 5	XXXIX
D.2.6	Other notes	XXXIX
D.2.7	Recommendations	XXXIX
D.3	Post-interview	XL
D.3.1	Questions and answers	XL
D.3.2	Recommendations	XLV
E	CityCop - Indentiface - Test Results	XLVI
E.1	Heuristic Evaluation	XLVI
E.1.1	Heuristics	XLVI
E.1.2	Evaluation	XLVII
E.1.3	Recommendations	XLVIII
E.2	Playtesting	XLIX
E.2.1	Test Participant 1	XLIX
E.2.2	Test Participant 2	XLIX
E.2.3	Test Participant 3	L
E.2.4	Test Participant 4	L
E.2.5	Test Participant 5	LI
E.3	Recommendations	LI
E.4	Post-interview	LII
E.4.1	Questions and answers	LII
E.5	Recommendations	LVI

List of Figures

1	Errors and failures discovered during the tutorial section of Age of Empires II	10
2	Start menu for "Alex and the Search for the Elemental Stone". . .	20
3	Introduction movie for "Alex and the Search for the Elemental Stone".	21
4	World Map	22
5	Inventory system	22
6	Start area of the Water level	22
7	The End level	23
8	Fancy Diving start menu and gameplay elements	27
9	CityCop Facesketch start menu and gameplay elements	30
10	CityCop Identiface start menu and gameplay elements	32

List of Tables

1	Methods Comparison	24
2	Methods Comparison	28
3	Methods Comparison	30
4	Methods Comparison	33
5	Methods Comparison	34
6	Heuristic rules for Gameplay	I
7	Heuristic rules for Game Story	II
8	Heuristic rules for Mechanics	II
9	Heuristic rules for Usability	II
10	Heuristic evaluation of Gameplay	III
11	Heuristic evaluation of Game Story	IV
12	Heuristic evaluation of Mechanics	IV
13	Heuristic evaluation of Usability	IV
14	Heuristic rules for Gameplay	XIX
15	Heuristic rules for Mechanics	XIX
16	Heuristic rules for Usability	XX
17	Heuristic evaluation of Gameplay	XX
18	Heuristic evaluation of Mechanics	XXI
19	Heuristic evaluation of Usability	XXI
20	Heuristic rules for Gameplay	XXXV
21	Heuristic rules for Mechanics	XXXV
22	Heuristic rules for Usability	XXXVI
23	Heuristic evaluation of Gameplay	XXXVI
24	Heuristic evaluation of Mechanics	XXXVII
25	Heuristic evaluation of Usability	XXXVII
26	Heuristic rules for Gameplay	XLVI
27	Heuristic rules for Mechanics	XLVI
28	Heuristic rules for Usability	XLVII
29	Heuristic evaluation of Gameplay	XLVII
30	Heuristic evaluation of Mechanics	XLVII
31	Heuristic evaluation of Usability	XLVIII

1 Introduction

This thesis looks at user-centered design methods for use in game development. Specifically, it tries to identify methods that are usable for small game development studios. There have been research efforts in creating methods that can be used for game development. The problem for small game development studios is that these methods often rely on a certain level of competence within the usability field. Since small game development studios consist of a small amount of people, it is unlikely that they have anyone with this level of competence. This can be a challenge for these developers. Additionally, several of these methods make use of professional testing facilities. This is not necessarily available to small developers, at least not without an additional cost. Some methods use large sample sizes. This requires organizing and takes up time and resources. This poses a challenge to small game developers as they are left without a structure designed for their needs. It makes it harder and more expensive to identify issues in their games.

The methods used in this thesis are meant to uncover issues in games. The finding of these issues are meant to guide game designers in their decision making. The methods are not game design methods as such. Nor are they meant to be used for scientific purposes.

Many methods have been developed for use in the software industry. These methods focus on usability and not on the total experience of the user. Games require a different approach. Traditional software produces the value outside of the software itself. Games are intrinsic in value and as such, certain elements that are important in traditional software are not important in games. There are significant differences between the approaches needed for the evaluation of games and traditional software. [Malone \(1982\)](#) refers to these two different categories of software as "toys" and "tools". Toys should have an uncertain outcome based on the user's input whereas tools should have consistent and reliable outcome based on the user's input. As games only have intrinsic value, the criteria for success are different.

[Federoff \(2002, p. 11\)](#) points out that usability language is typically not used in the game industry. This is another challenge. Methods need to be explained in a non-professional manner.

1.1 Problem description

The methods to be used by small game developers need to be usable on a small sample size. They must be easy to employ in a test. They must be understandable

for people not working in the usability field and they need to test the [playability](#) of a game rather than just the usability.

1.2 Document structure

After the introduction, there is a review of the existing literature on the subject of game testing methods and related subjects, this is placed in Chapter [2](#). The methodology for the study is outlined in Chapter [3](#). In Chapter [4](#) the test results are presented and analyzed. Chapter [5](#) presents a case study done with Suttung Digital, a small game developer studio. An interview is conducted in order to see how they have used the data provided to them from the test of their game *Fancy Diving*. Chapter [6](#) discusses the findings of the study and the limitations. Chapter [7](#) presents the conclusion and topics for further study.

2 Literature review

2.1 What is User-centered design?

It is generally considered that the field of Human-Computer Interaction, more commonly referred to as HCI, was formally founded in 1982 ([Lazar et al. 2010](#), p. 2). Although work had been done before this that could be considered within the realm of HCI. The field draws on experiences from a broad amount of expertise in other fields. Psychology, computer science, sociology, and industrial design are examples of already existing fields that have greatly influenced HCI.

HCI was in the early days an approach designed to measure performance in a quantitative way ([Lazar et al. 2010](#), p. 5). Measurements were taken on how many errors produced and how fast a user took to complete a set of tasks. These measurements are still in use, and they do serve a purpose when doing user tests on computer software. There are limitations to this approach. As an example, it is not possible to measure why someone does not use a particular feature of a software solution by measuring purely quantitative aspects. It became apparent that an approach that could answer questions like this was needed.

User-centered design methods came about as a result of this need. These methods try to answer questions surrounding the user and their needs. A term often used to describe the goal of this approach is *usability*. [Rubin et al. \(2011\)](#) outline the different criteria that are important to achieve this concept.

- *Usefulness* describes to what degree the product enables the user to achieve their goal
- *Efficiency* is how fast a task can be completed
- *Effectiveness* deals with the users expectancy of how the product operates
- *Learnability* is related to effectiveness and deals with how fast a user can learn it
- *Satisfaction* refers to how satisfied a user is when using the product and it therefore also includes the users expectations of the product
- *Accessibility* deals with how accessible the product is to the consumer base, important areas of interest are people with disabilities.

Not all of these criteria may be relevant to game development.

2.2 Why test games?

According to [Davis et al. \(2005\)](#), higher-quality games tend to sell better. This means that it is in the interest of the game developer to look for ways to improve their games. Improving games by looking at how their consumers approach their

products is a natural next step as it provides information on certain aspects of a game that would otherwise be overlooked. [Sánchez et al. \(2012, p. 1033\)](#) agrees with this and argues that having mechanics to test user experience in games is a key factor to ensure competitiveness in a market that is saturated with competing products.

It is clear that testing games is an important factor in ensuring that the quality of games is as high as possible in order to be competitive.

2.3 Why are games and game research important?

Games generate a significant amount of revenue ([Entertainment Software Association \(ESA\) 2015](#)). The total amount spent on video games in the by consumers in 2014 in the US was 15.4 billion dollars. If hardware and accessories used for gaming are added to this, the amount increases to 22.41 billion dollars. This means that games can represent a significant investment of both time and money. Not being able to make a return on this investment can impose large monetary losses on a game developer. This, in turn, means loss of jobs.

[Pagulayan et al. \(2007\)](#) argue that games is an important part of mainstream culture and therefore have a cultural and social impact on society. It also drives innovation in computer interface development as games can take more chances than traditional software. Games also push the limitations of computer hardware which drives the development of more powerful hardware. This is shown by [Shippy & Phipps \(2010\)](#) when they explain how the development of the *Xbox 360* and *PlayStation 3* needed new hardware solutions.

Keeping players motivated is a key factor in improving the user experience in games [Sánchez et al. \(2012, p. 1035\)](#). An important element for motivation is the level of fun that players are having while playing a game. In order to measure fun, we must first define it. Fun is an abstract concept, but there are elements that can be defined within this concept. Oxford dictionary defines fun as enjoyment, amusement, or light-hearted pleasure.

([Malone 1981, p. 335](#)) tries to identify different elements that create intrinsic value. According to him, intrinsic value is stronger when performing a learning task than the promise of extrinsic reward.

2.4 Does game testing differ from testing traditional software?

The difference between traditional software and games can be defined as the difference between tools and toys. Tools are used to achieve an external goal, whereas games have intrinsic value ([Malone 1982, p. 65](#)). The approach to testing should, therefore, take this into consideration. Traditional software should ensure a high productivity which comes from the software being designed with as few challenges as possible for the end-user. In contrast, games would be boring if there were no challenges. Boring games are bad games.

There are some aspects that are similar between the two. Games usually have

menus for changing options and functions for saving and loading saved games. Games may require that the user has a conceptual understanding of the content and the interfaces used in the game. These elements exist within the realm of traditional software as well. This means that usability tests can be run on both types of software with some success (Davis et al. 2005).

Usability testing can only test a limited part of the total user experience in games. One key criterion for the success of games is how much the player enjoys playing it. Fixing usability issues will help to achieve this goal, but other criteria need to be tested as well.

2.5 What can be tested in games?

Malone (1982, p. 65) developed the first known heuristics for evaluating games. The work was primarily focused on instructional games and not games designed mainly for entertainment. Three categories were created that contained testable heuristics:

- Challenge
- Fantasy
- Curiosity

The heuristics proposed show signs of being an early attempt at formalizing game evaluation and are interesting mainly for its historical significance rather than methods upon to build.

According to Clanton (1998), the elements of a game that are available for testing can be divided into three categories:

- Game Interface
- Game Mechanics
- Game Play

Game Interface includes game controllers and information displayed in games. *Game Mechanics* cover the physics and behavior of the game. *Game Play* involves the elements which give the player a goal in the game, like the narrative. These same categories were used by Federoff (2002) when compiling heuristics for a case study at a game developer.

Desurvire et al. (2004) used the following categories in their method called Heuristic Evaluation for Playability (HEP):

- Game Play
- Game Story
- Mechanics
- Usability

Game Play is the set of problems and challenges a user must face in order to win the game. *Game Story* includes all plot and character development. *Mechanics* involve the programming that provides the structure by which units interact with

the environment. *Usability* addresses the interface and encompasses the elements the user utilizes to interact with the game.

In a follow-up study to the development of HEP, [Desurvire & Wiberg \(2009\)](#) came up with a set of heuristics after discussions with several game developers such as THQ, Activision, Relic, Pandemic, Avalanche, Disney and Microsoft Game Studios. This method was named Game Usability Heuristics (PLAY). This method defines three categories:

- Game Play
- Coolness/Entertainment/Humor/Emotional Immersion
- Usability and Game Mechanics

There are similarities between the proposed testable elements. *Game Play* is present as its separate category in all three categorizations. *Game Interface* and *Usability* is the same thing in practice as they deal with how the user can interact with the game. One could argue that *Game Story* and *Game Play* could be considered the same category as the story in games is usually told through the game play, and there may, therefore, be benefits to collectively analyze them as one.

In PLAY, the authors decide to combine *Usability* and *Game Mechanics* into one category and the creation of a new category that deal exclusively with the emotional factors a player gets while playing a game. The authors explain this choice by stating that PLAY, unlike HEP, recognizes game design is not just an art but also a science. The potential problem with using heuristics for measuring emotional responses to a game is the large diversity of emotional responses that humans are capable of. The combined category of *Usability and Game Mechanics* could, and possibly should, be split into two different categories. Game mechanic issues are not necessarily the same as usability issues.

2.6 Can existing heuristics be used?

[Federoff \(2002\)](#) tested Nielsen's "Ten Usability Heuristics" to determine their usefulness in evaluating games. Federoff compared Nielsen's heuristics to established heuristics for game evaluation. The conclusion show that the heuristics could be used to evaluate game interface issues. However, it failed to find game-play issues.

[Pinelle et al. \(2008, p. 1454\)](#) argue that since these heuristics have been developed primarily for desktop applications, they have limited meaning in games. Certain things like camera angles in games and controller mappings will not be uncovered using these heuristics.

[Shneiderman \(1986, p. 60\)](#) created eight golden rules for design. These heuristics are analyzed in this thesis in order to see how applicable these rules are to games. Each rule is written out, and then a comment is made outlining how applicable they are to games.

1. Strive for consistency

"Consistent sequences of actions should be required in similar situations; identical terminology should be used in prompts, menus, and help screens; and consistent commands should be employed throughout."

Menu and graphical user interface items benefit from giving consistent feedback to the player. However, it may be desirable to have gameplay elements behave differently to increase the challenge level and variation of the game.

2. Enable frequent users to use shortcuts

"As the frequency of use increases, so do the user's desires to reduce the number of interactions and to increase the pace of interaction. Abbreviations, function keys, hidden commands, and macro facilities are very helpful to an expert user."

Although using function keys can be beneficial in certain complex games, usually large turn-based strategy games, there is little reason to decrease the amount of interactions in most games.

3. Offer informative feedback

"For every operator action, there should be some system feedback. For frequent and minor actions, the response can be modest, while for infrequent and major actions, the response should be more substantial."

Most games require an interface element which provides the player with necessary information. Elements such as life indicators or amount of ammunition left for a specific weapon.

4. Design dialog to yield closure

"Sequences of actions should be organized into groups with a beginning, middle, and end. The informative feedback at the completion of a group of actions gives the operators the satisfaction of accomplishment, a sense of relief, the signal to drop contingency plans and options from their minds, and an indication that the way is clear to prepare for the next group of actions."

The majority of games have some sort of dialog which gives closure to the player. This dialog often takes the form of a "Game Over"-screen.

5. Offer simple error handling

"As much as possible, design the system so the user cannot make a serious error. If an error is made, the system should be able to detect the error and offer simple, comprehensible mechanisms for handling the error."

In games, it is expected that the player will make errors at certain points when playing. These errors are often by design and as such it is not in the interest of the game designer to work towards an error free experience on the part of the player. Errors made when changing settings in the game menus should be easily handled.

6. Permit easy reversal of actions

"This feature relieves anxiety, since the user knows that errors can be undone; it thus encourages exploration of unfamiliar options. The units of reversibility may be a single action, a data entry, or a complete group of actions."

Many games have the ability to save and load game states. There are some games where this is not possible by design. Certain games do not allow for this and are designed in such a way that the player has to start again. Roguelike¹ games make use of this principle.

Roguelike games make use of procedurally generated dungeons in which the player faces off against monsters and various challenges. A new character must be created when the player dies. Death is permanent and save games do not exist.

7. Support internal locus of control

"Experienced operators strongly desire the sense that they are in charge of the system and that the system responds to their actions. Design the system to make users the initiators of actions rather than the responders."

In games, the players should not feel that they are in control of more than their representation in the game. If full control were given over all aspects of the game, it would no longer have any challenging aspects.

8. Reduce short-term memory load

"The limitation of human information processing in short-term memory requires that displays be kept simple, multiple page displays be consolidated, window-motion frequency be reduced, and sufficient training time be allotted for codes, mnemonics, and sequences of actions."

With the exception of memory games, it is not desirable that the player has to remember anything beyond the narrative elements.

Conclusion

The conclusion reached is the same as [Federoff \(2002\)](#) and [Pinelle et al. \(2008\)](#). There is some merit to identifying interface issues, but Shneiderman's heuristics are not suitable for finding issues with gameplay or game mechanics. Heuristics that take these elements into consideration are required in order to be able to use heuristics to an acceptable degree.

Attempts have been made to create heuristics specifically targeting games. An ambitious effort has been made by [Barwood & Falstein \(2006\)](#). Here they are trying to collect 400 heuristic rules that apply to computer games. 112 rules have been collected at the time of writing.

¹<https://en.wikipedia.org/wiki/Roguelike>

2.7 Which methods have been developed for game testing?

There have been several attempts at creating methods for testing computer games. [Federoff \(2002\)](#) developed a set of heuristics for game design by reviewing existing literature. The heuristics focus on storyline, gameplay issues, and engagement. The heuristics were refined after spending a week with a game developer where the heuristics used by the developer were identified. The findings closely matched the heuristics created from the literary review.

Rapid Iterative Testing and Evaluation (RITE) is a method developed by [Medlock et al. \(2002\)](#). The method focuses on quickly finding and fixing problems using iterative development. The most common use of iterative methods is called agile methods [Beyer \(2010\)](#). Two well-known methods are *SCRUM* and *XP (Extreme Programming)*. Both methods focus on developing software over what would traditionally be considered a short amount of time with a relatively small team of 5-7 people. Traditional software development would probably make use of more other methods similar to *RUP (Rational Unified Process)* or *Waterfall* ([Sommerville 2011](#)), though the latter is considered to be an outdated approach.

The game tested in the RITE-study is *Age of Empires II* ([Microsoft 1999](#)), specifically the tutorial section of the game. In total 15 participants played the tutorial. Changes to the game were made after test participants 1,2,5,8,9 and 10 (six iterations). The issues found were split into two categories; failures, and errors.

- Failures result in the inability to complete the tutorial.
- Errors cause confusion.

The focus on quick fixes and iterations means that newly implemented fixes can be verified as soon as they are done. The six last participants only encountered one issue, down from nine in the first test. There was a jump to 14 issues with the second participant, but there was a steady decrease of issues after the third iteration of the tutorial, as can be seen in [Figure 1](#).

In 2004 [Desurvire et al. \(2004\)](#) published a paper where they introduce a method they call *Heuristic Evaluation for Playability*, or *HEP* for short. This method focuses on using a set of heuristics in order to evaluate games. The results show that using heuristics in games is a viable solution.

[Desurvire & Wiberg \(2009\)](#) went on to create a new set of heuristics based on the success with *HEP*. This new method was called *Game Usability Heuristics*, or *PLAY*. The biggest difference from *HEP* was that this method recognizes games as both an art and a science. This method has a higher focus on trying to evaluate game content rather than just usability issues. The study conducted on this principle was considered a success.

The Playtest Method was developed by [Davis et al. \(2005\)](#). It uses an approach that draws inspiration from standard usability tests. It is designed to use a fairly

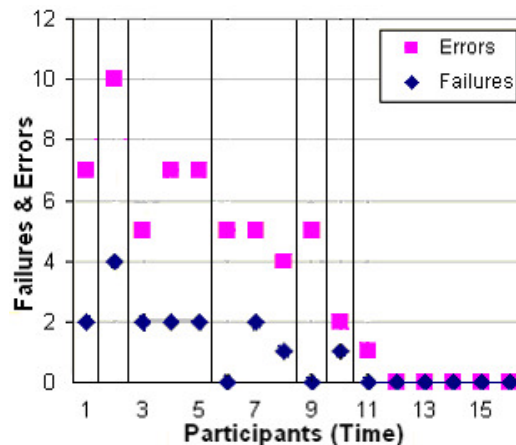


Figure 1: Errors and failures discovered during the tutorial section of Age of Empires II

large amount of participants, 25-35. The participants are given the installation media for the game and any documents that would follow the game upon release. Tests are run for an hour during which participants are observed. This is done to test the player's initial experience of the game. After playing, the participants are presented with a survey in order to uncover more information on how they perceive the game. This method was successfully used in order to test *Digital Anvil* (2003), a game published by Microsoft Game Studios for the XBOX.

Tracking Real-Time User Experience (TRUE) (Kim et al. 2008) was developed to try and explain why users made certain choices. The method focuses heavily on recording the users actions on camera and via recording software on the computer. A short survey is also conducted after each play session. Even though the method was not developed specifically for games, the case studies conducted on the methods employed games. Specifically, Halo 2 (Microsoft Game Studios 2004) and Shadowrun (FASA Interactive 2007) were tested. The *TRUE*-method was considered to be an invaluable tool that has been used by the authors to make improvements to more than 20 games in a variety of genres.

Pinelle et al. (2008) created a set of heuristics by identifying usability issues outlined in a selection of game reviews posted by *GameSpot*², an online video game review site. These issues were then categorized and grouped before the heuristics were developed. Five experienced evaluators were recruited in order to use these heuristics to test a game called *Necromania: Traps of Darkness*. After the evaluation, they were asked to outline the strengths and limitations of the heuristics evaluation. The end results showed that the evaluators understood and were successful in applying the heuristics.

In addition to these methods, there is the possibility of conducting Beta test-

²<http://www.gamespot.com>

ing. This method is fairly wide-spread and game developers regularly invite people to partake in Beta tests of their upcoming games. This method of testing has the potential of having a large number of participants. This is especially true when conducting so-called open beta tests, where there are no restrictions for entering. This method is good at finding bugs and other technical issues in games. It is not good for testing the gameplay experience of the participants. There is also a chance that the participants are the most advanced players and therefore do not represent the average gamer. Issues experienced as problematic for less experienced gamers may go unnoticed (Davis et al. 2005).

The methods outlined here show that there have been serious efforts in developing methods for testing games. It is also shown that there are limitations in using these methods for smaller game developers. The methods described assumes that the people conducting the tests are experts in the fields of conducting and evaluating user testing. For smaller game developers this can be an issue, as it is likely that they do not have anyone with this expertise. Hiring an expert in order to conduct the tests might also be unacceptable because of the potential cost associated with this. The test methodologies need to be broken down into more easy-to-handle methods.

2.8 Who should be recruited to participate in testing?

Medlock et al. (2002) used a group of participants that had never played a *Real Time Strategy* game, with a broad age range and both genders. All participants had played at least one PC game during the last year. Even with none of the participants having any experience with the genre a significant amount of issues were found and corrected.

It is the opinion of (Krug 2010, p. 41) that users with domain knowledge are not needed to uncover most usability issues. In fact, he argues that testing only on users with domain knowledge may cause problems for new users not necessarily familiar with the domain terms. Users unfamiliar with something may reveal things that users familiar with a domain won't reveal as they will not work around known issues.

This is in contrast to traditional usability testing. Lazar et al. (2010, p. 369) stress the importance of finding people that have the correct personal attributes and goals for your study. Domain expertise is stressed, something that is of less or no importance in games.

This shows that the approach traditional usability testing takes to test participants differ significantly from what is required for testing games.

2.9 How many test participants are needed?

Medlock et al. (2002) make use of a small amount of participants between implementing changes and verifying the success of these changes. During the user test on *Age of Empire II*'s tutorial changes were made after testing on only one

participant.

According to [Lewis \(1991, p. 3\)](#), it is the experience of many usability engineers that little new information is gathered by running user tests on more than five or six participants. This is also supported by an earlier study by [Lewis \(1990\)](#). The author states that in some cases an adequate amount of data can be gathered from as few as three test participants. Due to diminishing returns after the first participants, the number of test participants should not exceed ten participants.

[Krug \(2010\)](#) also claims that three users are enough to uncover the most significant problems. He also brings up a point of importance for small development studios. There is no point in uncovering more issues than you have the time or resources to fix.

It is important to be aware of the limitations of a small test sample. The discovery of an issue is dependent upon the frequency of occurrence within the user population ([Lewis 1990, p. 6](#)). This means that issues that have a low occurrence rate may go unnoticed when using a small user sample. However, the benefits of using a small sample size will in many cases outweigh the detriments as it is relatively cheap and efficient.

2.10 When should testing be conducted?

Steve Krug's recommendations in his book "Rocket Surgery made easy" ([Krug 2010](#)) is that testing should start as soon as possible. Krug argues that testing should be started even before there is much to show. His testing is done on the usability of websites, but the kind of issues that is related to a user's understanding of what to do remains the same as in games. Krug's recommendation of using one day every month for testing as a minimum is a reasonable approach that should be manageable by a small game developer studio.

2.11 When should testing stop?

It is likely that user testing will gather information on more issues than there are resources available to fix. This may potentially occur as soon as during the first user test. It is a good idea to stop testing when you can see that the issues found will take some time to fix. According to [Krug \(2010\)](#), it is a good idea to test once a month. Having an amount of issues to be fixed that will take about a month to implement is most likely a good approach.

2.12 What has been done to meet the need of small game developers?

As can be seen in Section 2.7, there has been little effort put into creating methodologies for small game development studios. This is probably due to much of the research being funded by larger game development studios, as is the case with [Kim et al. \(2008\)](#) and [Medlock et al. \(2002\)](#) which both did research for *Microsoft Game Studios*. This means that most of the test methods

developed require access to test facilities in order to conduct controlled tests. In all likelihood, smaller game development studios do not have access to these types of facilities.

Another issue is the required amount of test participants. The *RITE* method used 16 participants in total for their test of the tutorial system in *Age of Empires II* (Medlock et al. 2002). Initially, this may not seem like a lot of participants, but recruiting participants can take a significant amount of time. The *RITE* method also relies on a highly iterative development cycle. Not everyone is familiar or comfortable with this approach. Davis et al. (2005) use 25-35 participants in their *Playtest Method*. A number that is too large to be practically viable for a small game developer.

Heuristic evaluation methods for games have been developed, like the ones outlined by Desurvire et al. (2004) and Desurvire & Wiberg (2009). They created categories for evaluating different aspects of games. However, these methods rely on people who are experts in usability to conduct the evaluations. Small game developer studios are unlikely to have a usability expert on staff.

2.13 Which methods should be used by small game developers?

Qualitative research is much less time consuming and reveals more detailed information than quantitative research (Fulton & Medlock 2003). There is little point in gathering quantitative data as we are always looking for the answer to why something is done a certain way and not the answer to how many are doing it. Even so, qualitative testing requires the data to be analyzed. This means that the methods small game developers use must be easy to conduct with a small sample size, and the data must be easy to analyze.

Heuristic evaluations do not require any test participants and can be done at any time during the development process. A specific set of rules must be developed that takes into account that the person conducting the analysis is not a usability expert. This means that the rules must be descriptive, but not too complex.

Observation studies have been used in many of the earlier research studies into game testing. They are easy to conduct as the only thing needed is a device to play the game on or a paper prototype. A single participant is then observed while interacting with the game or prototype. Employed correctly, it can produce a significant amount of data.

Interviews also requires few resources. After a test participant has played a game, it is easy to conduct an interview in order to expand on the participant's experiences with a game.

2.14 Can the use of UCD-methods affect the game design intent?

The use of user-centered design methods is meant to help designers improve their product for the end-user. One concern with using these methods in games

is that the feedback might alter the original intent of the game. [Pagulayan et al. \(2007\)](#) show that using these methods can be to the benefit of the designer's original design vision. This was the case when testing *Halo* ([Microsoft Game Studios 2001](#)). User testing showed that novice gamers engaged enemies at a larger distance than the designers intended. Changes were made to alleviate this, and later user testing showed that they had been successful.

User-centered design methods may actually be preserving designers vision if correctly used.

2.15 Usability vs. playability

In usability testing, the tests are designed to uncover the following attributes ([Rubin et al. 2011](#), p. 4):

- usefulness
- efficiency
- effectiveness
- learnability
- satisfaction
- accessibility

These attributes focus on a person's need to accomplish a specific goal as efficiently as possible. In games the most important thing to look for is how much they enjoy using the product, [playability](#) is the name commonly used for this attribute. This attribute looks for how easy it is to play a game and how enjoyable it is to play.

2.16 How to rate issues

[Nielsen \(1995\)](#) proposes a system of five levels for rating the severity of usability issues.

1. I don't agree that this is a usability problem at all.
2. A cosmetic problem only: need to be fixed unless extra time is available on the project.
3. Minor usability problem: fixing this should be given low priority.
4. Major usability problem: important to fix, so should be given high priority.
5. Usability catastrophe: this is imperative to fix this before the product can be released.

Even though this serves to make it easier to prioritize fixing issues, this level of detailing the issues found during user testing is not recommended for someone that is not familiar with usability testing. It raises the required level of competence for analyzing user test data, something that is to be avoided in order to increase the ease of use of the methods.

2.17 Summary

There have been made several attempts to create and improve methods for game testing. Most of the methods outlined in this chapter require fairly extensive preparation and resources in order to be properly executed. A test facility is also required. The exception to this are the attempts at creating heuristics for game evaluation.

Heuristics are run without any test participants, but they require more expertise from the evaluator. If the heuristics are to work for small game developers, they need to be simplified so that this level of expertise is not needed.

Tests must be kept simple. If tests are straightforward, it is easier to get consistent results from session to session. All needs for specialized equipment and facilities must be reduced as much as possible.

As noted, the earlier the testing starts, the better. This means that even early prototypes made from paper can help to guide game design. This may be too much to ask from a small game developer studio, so a working and playable prototype solution is in all likelihood the way to go.

3 Methodology

3.1 Selection of methods

In this project, three distinct methods were selected for user testing. They are based on methods used in earlier research done in the field of user-centered design for video games, as outlined in Chapter 2. The three methods are variations of heuristic evaluation, observation studies, and interviews. These methods were selected as they fulfilled the criteria listed here:

- They can be conducted with little preparation.
- They are fast and easy to conduct.
- They can be used on small sample sizes.

All methods have been used in earlier studies and have produced good results when utilized. See Section 2.7 for more information on earlier developed methods.

As shown in Chapter 2, several different methodologies have been used in order to improve game testing. The selection of methods for small game developer studios needs to take into account that there may be no one that has usability experience.

The methods also have to work without having a test lab. Ideally, it should be possible to sit down in any location and run quick tests. In addition to this, the tests must provide reliable data on small test samples. Based on these criteria, the following methods have been selected:

Heuristic evaluation

Heuristic evaluations can be performed without any test participants. The challenge with this method is to come up with a set of heuristics that are not too complex for someone with little usability experience.

Observation/Playtesting

Observation studies are possible to conduct as long as there is access to a device on which the game is to be played. A test lab is not required, although it is preferable to do these tests in a relatively undisturbed area. For the purpose of game testing, this method is referred to as *playtesting*.

Interview

Interviewing a test participant after having played the game can give insights into why certain choices were made during gameplay. No specialized equipment or test labs are necessary to accomplish this.

3.2 Evaluating the methods

Several user tests were run to evaluate the efficiency of the proposed methods, and how they differ. Every test used the three proposed methods and they were conducted in the same order for every test. The order is as follows:

- Heuristic evaluation
- Observation Study/Playtest
- Interview

Heuristic evaluations provide test data without having to use test participants. The results of this evaluation are then used to improve the game before additional tests are needed. This ensures that errors are uncovered and addressed and not counted twice when other tests are run later on. The part of implementing changes before new tests are deliberately avoided as it is in the interest of the study to see how effective the tests are on their own. It would also have required a dedicated development team, something which was not available. However, during the test of *Fancy Diving* in Chapter 4.2, there were changes made to the game between the heuristic evaluation and the observation study.

The heuristics used are described in each test as there are variations. Some changes to the heuristics are refinement based on earlier tests and some are changes based on the game that was tested. Not all heuristics apply to all types of games.

A playtest was conducted after the heuristic evaluation. The amount of participants vary between the tests, and this is described in the sections containing the specific tests. There was also some variation in how the tests were conducted. There were certain aspects of the tests which did not make sense for all types of games. This variation is described in the specific sections containing the test.

An interview was conducted immediately after the observation study has finished. This is a logical approach as the test participant would have needed to play the game first. Using the same test participant also shortens the recruitment need, and it is also the way it is intended that small game developers would use the methods proposed. This test is meant to expand upon findings during the playtest.

3.3 Selection of participants

There are no criteria set for the recruitment of test participants. This is to see if valuable data can be generated when recruiting random people. Lowering the recruitment standard makes it easier for small game development studios to recruit people for testing. This corresponds to [Krug \(2010\)](#)'s suggestion as outlined in Section 2.8.

3.4 Results

Results are based on the amount of issues uncovered during the different tests. An issue is anything that negatively affect the gameplay experience. The results are open to interpretation, as all the data gathered are qualitative.

Personal experience and commentary on how well the different methodologies work are also added as part of the result of the test results.

3.5 Analysis and evaluation

The success of a test is based on the amount of issues found and on personal experience with the different methods. During the analysis, the amount of issues found using each method is quantified to give insight into how effective the methods are.

4 Results and analysis

In this chapter the different methods outlined in Section 3.1 are used to test a selection of four different games. The purpose of these tests is to identify how the methods differ and how to best apply them. At the end of each test, the methods are evaluated separately. After this, the methods are compared to see where and how they differ in practice.

4.1 First Test - "Alex and the Search for the Elemental Stone"

4.1.1 Test description

This is done by using methodologies to analyze a specific game, "Alex and the Search for the Elemental Stone" (Lieng et al. 2013). This game was created for the final delivery in a programming course as part of a bachelor's program. It has known issues and bugs and is selected for this specific reason.

"Alex and the Search for the Elemental Stone" is a point-and-click adventure game. The game draws upon influences from old games in this genre, such as "The Secret of Monkey Island" (LucasArts 1990) and "Day of the Tentacle" (LucasArts 1993). The game puts the player into the role of Alex. A wizard's apprentice which goes on a quest for the Elemental Stone after burning down his master's tower.

This chapter will begin with a section containing an overview of the game. The section after this will compare the results from the methodologies. The last section will look at what the strong and weak points of the methodologies are and make recommendations for how to improve upon them. For the complete test results and more detailed description of the methodologies used, see Appendix A.

Heuristics

The heuristics used are derived from the work of Desurvire et al. (2004) and Desurvire & Wiberg (2009). Heuristics that are not applicable to the selected game have been excluded from the heuristic evaluation. The game is first played, and then the heuristics are applied. How well the game conforms to a selected heuristic is identified during this process. Suggestions are written that explain what changes are necessary to make the game conform to said heuristic. Four categories of heuristics are used; Gameplay, Game Story, Mechanics and Usability.

An example of a heuristic is "Provide immediate feedback for user actions".

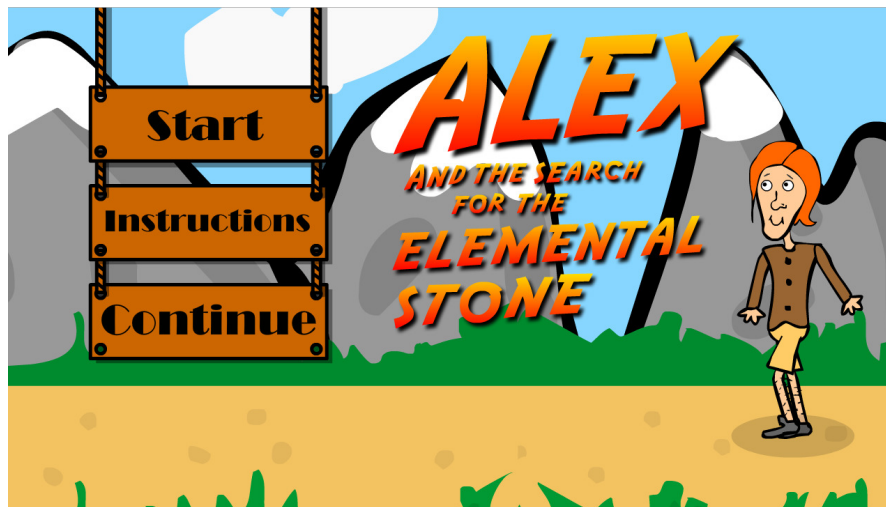


Figure 2: Start menu for "Alex and the Search for the Elemental Stone".

Playtesting

During the playtest, test participants are observed, and notes are made of issues and bugs that show up during gameplay. The test participants are asked to talk out loud when a problem is encountered, or there are other issues that they do not understand. Participants are told to play until they complete the game, or they do not feel like playing the game anymore, for whatever reason. If they quit before the game is completed, they will be asked what caused them to end the game. If the test participants have any questions about the game, they are not to be answered until the playtest is over. This is made clear before the playtest starts.

The notes taken are written into longer narratives after completion of the playtests. This is done to provide context for anyone not participating in observing the playtests, which is how it is meant to be done when a developer employs the method.

Interview

After the playtest, the interview is conducted. The questions asked are based on the heuristics used in phase 1. The approach for the interview is a semi-open structure. The questions for the test participant are asked as they are written by the interviewer, but follow-up questions are encouraged to expand upon the answers given.

4.1.2 The Game: "Alex and the Search for the Elemental Stone"

Figure 2 shows the start menu of the game. Once the player hits the "Start"-button an introduction movie is played, as seen in Figure 3. The "Instructions"-button displays information on how to play the game and the "Continue"-button



Figure 3: Introduction movie for "Alex and the Search for the Elemental Stone".

allows the player to select which areas have been completed earlier.

The first interactive part of the game presented to the player is an overview of the world, as seen in Figure 4. The world is split into five different levels; Fire in the lower left part, Earth in the upper right, Air in the upper left, Water in the lower right and the End level which is represented by a mountain in the middle of the map. The End level is not selectable until the other four levels have been completed.

When a level has been selected, the player starts in an area of this level. The player is then able to move around using a mouse as a controller. There are several areas within each level, and the player can freely move between these areas. Certain items have been placed around the areas, and the player is able to pick these up. Once an item has been picked up, it is placed in the inventory which is available in the lower right corner of the screen, as seen in Figure 5. The items that have been picked up can be used to interact with objects in the game world. A short movie is played once an item is used on a correct object.

The start area of the water level can be seen in Figure 6.

A "mini-game" must be completed when the player has solved all tasks in a level. This "mini-game" takes a different form for all the levels. The Fire level has a three-in-a-row game, the Earth level has a memory game, the Water level has a labyrinth game and the Air level has a shooting game.

The End level is a side-scrolling platform game where the player must jump over obstacles and avoid fireballs that appear at random. See Figure 7 for an example of how this level looks. The game is finished once the player makes it to the end of this level.



Figure 4: World Map



Figure 5: Inventory system



Figure 6: Start area of the Water level



Figure 7: The End level

4.1.3 Comparison of results

Methods Comparison

Heuristic Evaluation issues

- Gameplay: 5
- Game Story: 3
- Mechanics: 1
- Usability: 2
- Total: 12

Playtest

- Issues: 19
- Bugs: 2
- Features: 2

Interviews Results - Issues

- Gameplay: 1
- Game Story: 2
- Mechanics: 1
- Usability: 3
- Total: 7
- Bugs: 1
- Features: 1

Interviews Total

- Gameplay: 4
- Game Story: 4

- Mechanics: 1
- Usability: 7
- Total number of issues: 16
- Bugs: 1

Table 1: Methods Comparison

Methodology	Issues found
Heuristic evaluation	12
Playtesting	23
Interviews	17

The Playtest uncovered the largest amount of issues and bugs in the game. [Desurvire et al. \(2004\)](#) found more issues using heuristic evaluation than they did during the playtests.

Even though the interviews did manage to uncover almost as many issues as the Playtest, the information gathered does not contain the same complexity. This makes it harder to come up with viable solutions to improve said issues. Basing the interview questions directly on the heuristics makes it easy to come up with questions. However, this also limits the amount of new information that can be gathered after the playtesting. On its own, without an observation study, this can be beneficial. However, when also conducting a playtest and a heuristic evaluation beforehand, the questions could benefit from being more tailored to a specific game. Elaborating on the issues found in the heuristic evaluation rather than emulating the heuristics through a question might prove more fruitful.

4.1.4 Evaluation of methodologies

Heuristic Evaluation

The heuristic evaluation did find some major issues with the game. The biggest challenge with this approach is finding the right heuristic rules. When done correctly, major issues can be found with relative ease using this approach. Fixing these issues before the playtest is likely to reduce the amount of issues found in later stages of the game testing, as many of the same issues were also found during the playtest and the interviews. This will either reduce the amount of issues found or uncover other issues that may have been hidden in the later stages. Both outcomes are beneficial to the game development process. The lack of gameplay feedback is an important problem with this approach, as it is not possible to uncover how enjoyable the game itself is with a target audience. Certain types of issues may also be out of reach of a heuristic approach. Issues specifically related to how enjoyable the audience finds the game is one example of this.

Playtesting

The playtesting did find the largest amount of issues, but the downside of this methodology is that it does not give much insight into why a player had a specific issue or why certain choices were made. One way to resolve this problem is to have a more interactive playtest, where the observer asks questions while the test participant is playing. A potential issue with this is that it may interfere with the how the test participant would normally play the game. This can lead to skewed test results.

Post-interview

The interviews found almost as many issues as the playtest. However, the approach used limits the elaboration on issues found during the playtest.

Suggestions for further testing

The heuristic evaluation needs to use heuristics that are relevant for the game being tested. New heuristics not derived from earlier work will be developed and employed to check how efficient and challenging this is.

The playtest worked quite well and was the best at finding issues with the game. No suggestions for change are necessary at this stage.

The interviews should not be as closely based on the heuristic rules used for the evaluation. In all likelihood, many of the issues found in the evaluation may already have been solved before the playtest. For future use, the methodology will be based on observations made during the playtest to elaborate on what has been observed. Some more generic questions will also be used, such as "How did you find the gaming experience?", as they can answer questions the playtest cannot answer.

4.1.5 Other Notes

Short notes were written during the playtesting. These notes were written into a longer narrative to give more context to other people reading the text, helping them to understand the player's context and motivation.

The time total for each participant was about an hour. Of this time, the playtest took roughly 30-40 minutes, and the interview took roughly 20 minutes per participant. Having no time limit for the playtest posed no problems. There may be a need for having a set time limit when testing larger games or when having tests run in quick succession.

4.2 Second Test - "Fancy Diving"

4.2.1 Test description

Between the heuristic evaluation and the playtests, the game was updated. Some of the changes made rectified issues found during the heuristic evaluation. This was somewhat helpful as changes between the heuristic evaluation and the playtest are intended to happen. The heuristic evaluation was performed on "Fancy Diving" version 0.1 #17 while the rest of the study was done on version 0.1 #18.

See Appendix B for the complete overview of the test results and methodologies discussed in this chapter.

Research Goals

The series of tests that were performed on *Fancy Diving* is performed with the intention of seeing if the methods that have been further developed works better than those used in Chapter 4.1. Since the game used is not the same, a direct comparison is not possible. The evaluation of the methods is based in part on the perception of how the methods worked in practice.

Heuristics

The heuristics used in Chapter 4.1 were split into categories. This categorical division will be reused with some modifications. The story category will be removed as the game has no story of note.

Playtest

Test participants were observed while playing the game. All participants were asked to talk out loud during the playtests. Issues they encounter during game play were written down. As it was not possible to complete the game, test participants were asked to play for 20 minutes each or until they did not want to play anymore.

Interview

A set of predefined questions and questions based on observations during the playtests were used to uncover potential issues and to illuminate issues that occurred during the playtest.

The interview questions are designed to be open-ended questions (Lazar et al. 2010, p. 111) when applicable. This forces the interview subject to consider the answer before giving it. Unlike the questions used in Section 4.1, these questions are not grouped into categories.

At the time of the user testing, a substantial redesign of the trick phase was being worked on. After the first round of questions, the current state of this new design was presented to the test participant. They were then questioned about this new design.

The redesign written by Suttung Digital is available in Appendix C.

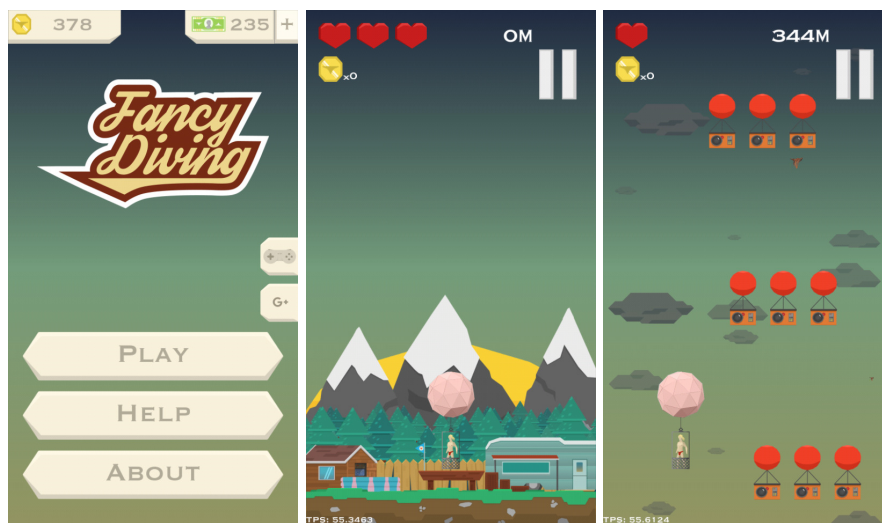


Figure 8: Fancy Diving start menu and gameplay elements

4.2.2 The Game: "Fancy Diving"

Fancy Diving is a game currently under development by Suttung Digital. The target platform is mobile devices running the Android operating system.

In the game, the player controls a character as he uses a hot air balloon to reach as high an altitude as possible. During this ascent, the player must avoid objects that will take health away from the player. Once all the health is gone the player character dives out of the hot air balloon. During this descent, the player continues to avoid obstacles. They no longer take away life, but they stun the character for a short time. The goal during this is to pick up coins.

Once the player reaches the bottom, he gets bounced back up again. The character now has to perform a series of tricks that will earn him currency when correctly performed.

4.2.3 Comparison of results

All three methodologies worked well in discovering issues that negatively impact the game. Most issues were discovered regardless of the methodology used. The amount of issues found while using the different methodologies can be seen in Table 2. This shows that there are only minor differences between the methodologies used. All are close to being equally effective when employed. The fact that some of the issues found during the heuristic evaluation were solved before the user testing got started shows that this type of evaluation can be used to find real issues.

The interview part of the test gave the best insight into changes beyond just the negatively impacting issues. Using this method, it was possible to get information on what the participants would like to see as new features in the game.

This includes suggestions for what type of upgrades that should be available for purchase in the game. This approach can be used to assist creative issues during game development.

Methods Comparison

Table 2: Methods Comparison

Methodology	Issues found
Heuristic evaluation	12
Playtesting	14
Interviews	12

4.3 Evaluation of methodologies

Heuristic Evaluation

The heuristic evaluation worked well in uncovering several important issues. The changes made from the ones used during the evaluation of "Alex and the Search for the Elemental Stone" worked well. Employing the heuristics to the game met with little challenge.

Playtesting

Playtesting *Fancy Diving* worked well in uncovering issues with the game. There is a potential for some issues when observing someone that is playing on a mobile phone, as it moves around more than stationary systems would. This problem did not occur during this playtest.

No players reached the set playtesting limit of 20 minutes as all participants felt that they were done with the game before this. The participants played for this amount of time:

- Participant 1: 16 minutes
- Participant 2: 8 minutes
- Participant 3: 8 minutes
- Participant 4: 6 minutes
- Participant 5: 11 minutes

Post-interview

During the post-interview, several new issues were uncovered. Several issues uncovered during the playtesting and heuristic evaluation were repeated.

4.4 Third Test - "CityCop Facesketch"

4.4.1 Test Description

CityCop Facesketch is a less complex game than those tested previously. The game was tested with a heuristic evaluation to determine how effective it would be on less complex games. A playtest was then performed on five test participants.

See Appendix D for the complete test results and more details on how the tests were done, including heuristics and questionnaire.

Research goals

The goal of this test is to determine how well the different methodologies work when used for a less complex game.

Heuristics

The heuristics used are the same as those previously used in Chapter 4.2 for the heuristic evaluation of *Fancy Diving*.

Playtesting

Test participants were put in front of a laptop and told to start the game when they were ready. Participants were asked to talk out loud, explaining what they were doing and why. An observer made notes on comments and what they did in the game. Participants were asked to play the game for as long as they wanted.

Interview

After the playtesting the participants were asked questions. These questions were general in nature and focused on the general experience of the game. No game specific questions were prepared for this test.

4.4.2 The Game: "CityCop Facesketch"

In *CityCop: Facesketch* the goal is to recreate a face with available graphical facial features. The game is played through a web-browser, but the final product will be playable on other devices, like tablets. It is part of a larger group of games that will be made for the CityCop-project. This is a transnational project that aim to improve security and safety in neighborhoods. More information can be found on <http://citycop.eu/>.

Screenshots from the game can be seen in Figure 9.

4.4.3 Comparison of results

There was little variation in the amount of issues found between the three methods used, as seen in Table 3. However, while using the different methods, the interview did have an edge in finding issues. If game specifics questions had been added it would likely have produced even more issues. Most issues were found regardless of the method used.

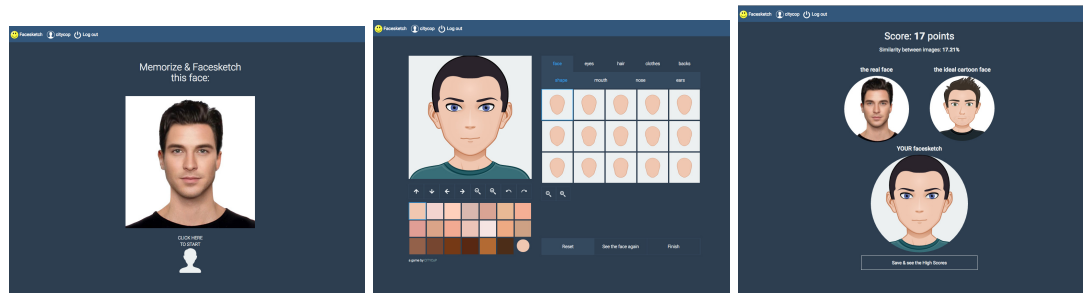


Figure 9: CityCop Facesketch start menu and gameplay elements

Methods Comparison

Table 3: Methods Comparison

Methodology	Issues found
Heuristic evaluation	7
Playtesting	6
Interviews	8

4.4.4 Evaluation of methodologies

Heuristic evaluation

Although the heuristic evaluation did find an amount of issues comparable to the other methods, using it did not feel effective. There are severe limitations to finding issues that affect the enjoyment of the game.

Playtesting

The game is not complex and very straight forward in nature. Just observing the participants would have been unproductive. Issues found with this method came from the users when they were talking out loud as they had been asked to do.

Interview

Interviewing the participants did produce the largest amount of issues, and it also felt like the method that worked the best. It was easier to produce results with this method compared to the other methods used.

4.5 Fourth Test - "CityCop Identiface"

4.5.1 Test Description

CityCop Identiface is a less complex game than *CityCop Facesketch* that was tested in Chapter 4.4, which was not a complex game itself. It is more conceptual in nature, and this test and the test results reflect this. The main purpose of this test is to test features and functionality in order to determine which will work the best and thereby deciding what to use when going forward with the development of the game.

There are three versions of the game, all with different variations of the same game. All three versions are tested to see what features work the best. The developer of the game had a list of specifics that he requested to have tested:

- Time limit or not?
- 4 or 6 faces (or more)?
- Is interaction easy and user-friendly?
- Tablet or Smartphone?
- Portrait or Landscape tablet/smartphone orientation?
- Good/decent gaming experience, fun?

Questions for the post-interview were based on this request. Tablet or smartphone was not tested due to the limitation of the available hardware. The game was made for the Android operating system, and an Android-based tablet was not available during testing.

See Appendix E for the complete test results and more details on how the tests were done, including heuristics and questionnaire.

Research goals

The goal of this test is to see how well methods employed work on games that are in a conceptual state where specific questions need answers in order to guide the development direction.

Heuristics

The heuristics used are the same as those used in Chapter 4.2 for testing *Fancy Diving* and Chapter 4.4 for testing *CityCop Facesketch*.

Playtesting

Test participants were given a mobile phone and asked to play the three different games in order. They were asked to talk out loud about what they were doing and why they were doing it. This was then observed. Notes were written while observing, one specific thing that was noted was how well they performed in finding the correct faces.

When they had played through all the version of the game they were asked to stop in order for an interview to be conducted.

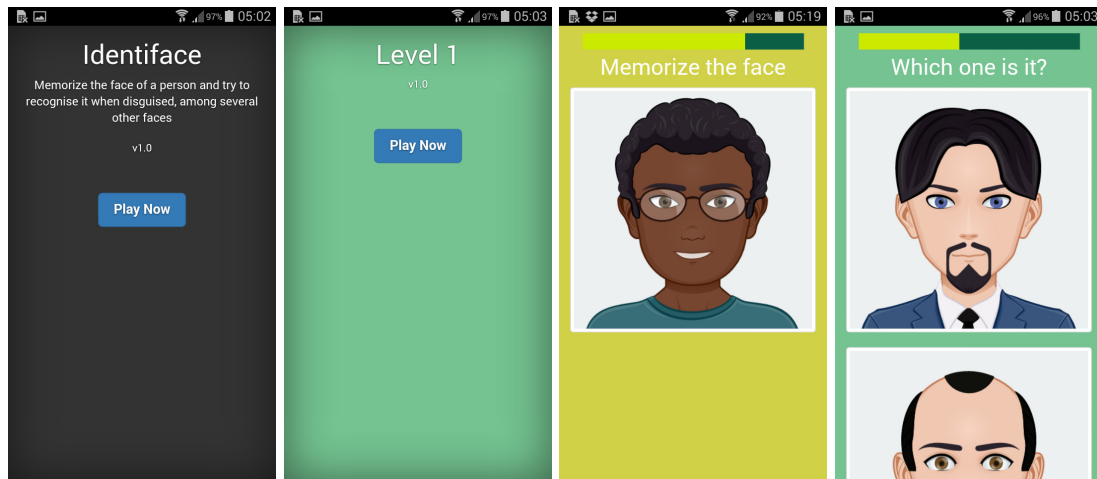


Figure 10: CityCop Identiface start menu and gameplay elements

Interview

The questionnaire used had some general questions that had also been used in earlier tests. In addition to this, some specific questions were asked that aimed to answer the questions posed by the developer, as outlined in Section 4.5.1.

4.5.2 The Game: "CityCop Identiface"

In *CityCop Identiface* the player is presented with a face. This face must then be found in a group of other faces. Certain features may have been changed in order to create a disguise for this face. This includes elements like hair style and color, wearing glasses, and other facial features. The player either scrolls or swipes through the faces and selects the one believed to be the correct face. Information on whether this was correct or not is then displayed. A score is presented if the selection was the correct one. As mentioned with *CityCop Facesketch* in Section 4.4.2, this game is also part of the *CityCop*-project.

The three versions of the game differ in the following way:

1.0:

- Level 1 & 2: Tapping – 4 faces – 10 seconds to pick
- Level 3 & 4: Tapping – 6 faces – 10 seconds to pick

v1.1:

- Level 1 & 2: Tapping – 4 faces – No time limit
- Level 3 & 4: Tapping – 6 faces – No time limit

v1.2

- Level 1 & 2: Swipe – 6 faces – No time limit

Screenshots from the game can be seen in Figure 10.

4.5.3 Results

Methods Comparison

Table 4: Methods Comparison

Methodology	Issues found
Heuristic evaluation	5
Playtesting	4
Interviews	8

4.5.4 Evaluation of methodologies

Heuristic evaluation

The heuristic evaluation did find some gameplay issues. It is not a good method for answering the specific questions stated by the developer.

Playtesting

Due to the simple nature of the game, observing the player did not produce much data. In trying to compensate for this, the correct and incorrect choices made were noted. This did reveal that the game is most likely too easy due to the high amount of correct answers.

Since the game is so easy and simple, talking out loud only produced a small amount of data. These data were of value, so there is some credence to conduct a playtest.

Interview

Interviewing participants found the largest amount of issues. It was also the method that felt the easiest to conduct. It also showed itself as being good at getting answers to specific questions.

4.6 Analysis of test results

The total amount of issues found during the tests are seen in Table 5. When quantified it appears that the playtesting and interviews finds more issues than the heuristic evaluation. The tests conducted does not take into consideration that found issues are meant to be addressed before conducting the playtest and interview. There was some changes made during the testing of *Fancy Diving*, but these were not significant enough to have an impact on the end result.

Table 5: Methods Comparison

Methodology	Issues found
Heuristic evaluation	36
Playtesting	47
Interviews	45

When analyzing the methods used in a non-quantifiable way, the results are more interesting. The analysis of the individual results after each test reveals that some methods are better suited for some purposes. Games that are simple and conceptual did not work as well with heuristic evaluation. There were also issues while playtesting. Using interviews was consistently the method that worked well in all scenarios.

5 Case Study - Interview with Suttung Digital

An interview with *Suttung Digital*¹ was conducted to get an impression on how a small game developer studio would approach user-centered methods. They had been provided with the results obtained from testing their game, *Fancy Diving*, and my analysis and recommendations. A meeting had been held when the analysis was completed, and the results and recommendations were presented to the development team. The interview took place several weeks after this meeting. This was deliberately done to give Suttung Digital time to start implementing changes based on the recommendations. Four members of *Suttung Digital* attended the interview.

The goal of the interview was to get a discussion going to try and explore their impression on user-centered methods and what they got out of the test report. A list of questions had been prepared to guide the discussion through different topics.

The interview is reproduced in its entirety in the following section. This is followed by a summary of the important parts of the interview. After the summary, there is a section that described the outcome of the interview.

5.1 The interview

1. What benefit did you get from the results of the conducted evaluation and user testing?

Suttung Digital: We got quite a bit from the test report. We were aware of many of the issues. It served as a confirmation on what direction we should take and what we have to do. It is good to get this information from other people as well. It is easy to sit in a bubble and only do the changes that we might like, but that others might not like. The test was done when the game was still early in development, so many of the experimental features had yet to be implemented. There were some things that we did not think about. Like a lack of gameplay variation. We were somewhat aware of this, but it is still good to hear someone else confirm this as you work harder to figure out what we can do to increase the variation. One of the things we got out of the test report is to create more variation in the hazards. This has become more of a core thing in the game that I thought it would be in the beginning. During the first two weeks after the report, we designed two new hazards.

Interviewer: Like the birds?

¹<http://suttungdigital.com/>

Suttung Digital: Yes. We have also added some mini-boss encounters. Like a ship that shows up and shoots like a bullet-hell pattern type game. When you get closer to space, you will meet a Gundam-mech that you must avoid charges from. We have had no inhibitions when it comes to this. We have managed to design two new encounters each week. We have a large backlog with many creative ideas that we can make use of.

Interviewer: So the feedback got you to think in other directions?

Suttung Digital: Yes, basically. We felt that the first phase of the game was the most natural place to create variation in gameplay. The second phase is very fast paced and somewhat shorter. Phase one could be extended. When the tests were conducted, we were still trying to decide what the core mechanics of the game were supposed to be. Another useful part of the feedback was the trick sequence in phase three. When we initially came up with this, we thought it would be very natural and work really well. Even though we said during development that a touch-based joystick never works, we still ended up doing it. We got that it really did not work when we tested it ourselves.

Interviewer: Yes, none of the test participants understood what to do. I managed to understand it, but I tried it for quite some time.

Suttung Digital: And when you understand how it works, it is still not good. Doing "Street Fighter"-combos on a touch screen won't work.

Interviewer: What did you end up with? Did you go for the swipe mechanic?

Suttung Digital: Yes. Large arrows appear that show the way you need to swipe. This needs to be done within a set time. A timer shows how much time you got. The right move needs to be performed before the timer expires. A new trick shows up when a trick succeeds, fails, or the timer expires. *At this point, one of the developers starts the game on his phone to show how the game works at the moment.*

Interviewer: Do you only have to make one swipe per trick?

Suttung Digital: Yes. A random direction is shown. There are eight different directions.

Interviewer: I can see that you have also moved the character up a bit.

Suttung Digital: Yes. We moved the character up as suggested. You no longer cover the character with your finger. We also moved the health indicators down. The idea behind this is that it is closer to the balloon and therefore easier to see when you lose life. From an art perspective, it was not as nice with different elements spread out across the screen. Placing the elements at the bottom of the screen made the UI more cohesive.

Interviewer: Yes, the UI looks more complete now.

Suttung Digital: The only problem is that on some button-less phones you get this *(shows an Android menu that shows up when the finger gets too close to the bottom)*. Not sure how to solve this issue as of yet. You can also see that we have added text that says "Press to start" in the beginning. We have also played with

the idea of adding a combo system.

Interviewer: Like doing several swipes within a set amount of time?

Suttung Digital: Yes, and something like *Gears of War* has when you reload a gun. A short interval when the swipes must be done to get a higher score. We have not completed this design yet. Another thing you mentioned was that touch worked well in the first phase, but tilt worked better in phase two. So we are looking into implementing options that allow the user to select which control he would like for the different phases. This is not something you often see, that you can select different controllers within the same game. We have also added a height high score.

Interviewer: Have you removed the height limitation?

Suttung Digital: We have added a line that indicates how high you reached during your last game. We have added more balloons. The different balloons have different height limitations. In the version you tested, there was a limitation of 800 meters. Now the max height is set to 550 meters. We have made an adjustment to this based on feedback and internal discussions. We also looked at how long we want the game to last, considering that it is a mobile game. It cannot last for too long.

Interviewer: So you have gotten something from the test results?

Suttung Digital: Definitely.

2. Do you have an approach that you use to uncover issues with Fancy Diver?

Suttung Digital: We do testing after each sprint. We have a session where we test the game individually. We look for things that have been documented during the last sprint. We make notes of what we find to be issues with the game. Then we have a discussion on what to add to the backlog. New features or changes, things like that. We only do internal testing.

Interviewer: So you do not bring in anyone for testing?

Suttung Digital: Some master student in my class play the game occasionally and provide feedback. Nothing structured.

3. Are the methods that I have used something you would be willing and able to use?

Suttung Digital: I think we would be able to use them.

Interviewer: Do you have any plans for using them?

Suttung Digital: We have no concrete plans for this, yet. It would be natural to do some testing closer to release. When the game is more cohesive. It will be easier to see where the problems are when we have implemented some more features. Maybe get some external test participants and run some tests. Very nice to have a framework available for this, as we do not have to research it ourselves.

Interviewer: Heuristics are often meant for people with some competence within areas like usability. Do you think that you would be able to use the ones that I have used? Or is it too advanced?

Suttung Digital: Many of these rules are things that I am looking for continuously. I am aware of them. It might be good to have a list of rules.

Interviewer: One of the benefits with heuristics is that you do not need to get people for testing. My thought on the test process is that you start with a heuristic evaluation, and then improve the issues found before running tests on users. This can eliminate double findings of the same issues.

Suttung Digital: It would be very useful when used like that.

Interviewer: I have tried to make the heuristics more direct and easier than they often are.

Suttung Digital: Yes, I feel that they are pretty clear. The heuristics are absolutely something we can use. They are concrete and easy to understand. As previously said, many of them are things that I am conscious about. A complete list is good to have if we were to have a process where we go through all then heuristics and make an assessment on each one.

Interviewer: During my testing I found that many of the issues found during the heuristic evaluation popped up again during the user testing. If they had been fixed before the user testing, we might have ended up with different results.

Suttung Digital: We had some of the heuristics in the backlog or were consciously aware of them.

4. Do you see any challenges with using the methods?

Suttung Digital: Getting people for testing might be a challenge, but it is not a major one. Time consumption and resources are the two major challenges as we are only four people.

Interviewer: Would you have tried to recruit from NTNU Gjøvik?

Suttung Digital: When we did some testing on a prototype for an earlier game, we used the bachelor group for which we were the principal. So the college is definitely a resource. We have discussed using the game development course. We would probably test on friends first. It is the easiest, but not necessarily the best option. We have some thoughts on the target audience, it gets clearer as we approach release.

5. How often would you be able to run tests?

Suttung Digital: That depends somewhat on what we want to test. We will probably not be able to do it every sprint, but it is not likely to be needed that often. Between milestones perhaps.

Interviewer: How often do you have milestones?

Suttung Digital: It is not set. We know somewhat internally when we reach them. Probably one time every month or so. Running tests every week is too much.

Interviewer: And there is no point in finding more issues than you can fix. Running tests too often may not give any benefit.

Suttung Digital: We do not implement so many features that there is any

point in testing it very often.

6. Which approach do you use when developing games? (SCRUM/XP/RUP)

Suttung Digital: SCRUM. We have weekly sprints. We used to have two-week sprints, but that was too long. We sometimes find things during a sprint that does not work as well as we thought. If we are locked into two weeks, this can become a problem. It was not flexible enough. When we are as small a team as four persons we have the option of being very agile, so it is good to exploit this.

Interviewer: Do you focus on features or bugs?

Suttung Digital: Features and tweaks, but mostly features.

Interviewer: Do you often introduce bugs when you implement features?

Suttung Digital: We have a process where we peer-review the code. We also test that the new feature works as intended. We also try to make sure that the rest continue to work, but this gets harder as the code gets more complex. It is too much to test every time. We do some testing on different OSes, and some things that work in Linux might not work in Windows. Our process helps, but some bugs still manage to sneak in. If we get breaking features, we try to fix it as soon as possible. Other bugs we put into the backlog.

7. How long would it take to implement the changes recommended in the test report?

Suttung Digital: Some of the findings are very small and easy to implement, like the "Touch to start" text at the beginning of a game. Changes to the control system have taken a couple of weeks. Changes are made in parallel with other work on the game. Getting more variation, like the hazards, is done over a longer period of time. Hazards have become more complex. Cosmetic upgrades are things that we have planned. We have to make the assets first and then the system around it. Costs and the like. This is not the highest priority, though. The game has to be cool first, preferably.

Interviewer: So it might take a couple of months in total?

Suttung Digital: Yes, with absolutely everything it could take up to three months I think.

Interviewer: What if we disregard the cosmetic features? Only features and bug fixes.

Suttung Digital: Hazards take quite some time, as it needs balancing. The in-game tutorial takes a couple of weeks. It has to be designed and then tweaked. It needs a bit of polish as it is the first thing you see in the game.

8. Are the methods understandable?

Suttung Digital: Yes, I think so.

Interviewer: What looks like the most challenging part?

Suttung Digital: I would probably have to have a couple of user tests before I can perform the tests in a consistent manner so that I do not confuse the user

because I am a bit unsure of how to do it.

9. Do you think that you would have analyzed the results the same way I did? (Same conclusions and recommendations)

Suttung Digital: I think I would have reached similar conclusions. We would have looked at what people had problems with and then we would have had a discussion afterward, where we would have tried to come up with solutions. That is how we usually do it. We did not agree with all of the recommendations. One of the things mentioned was that the meteor should do more damage, but this was in contrast to our design philosophy. Everything only does one point of damage. We try to keep our artistic vision. We do not want to make changes just because something is being said. What we want to do with the game also matters.

10. Would you be able to adapt the methods to your needs if needed?

Suttung Digital: I think so. We know how we want the features to work for the users. So it is a question of observing if they play the game as we intended it to be played.

Interviewer: Like asking follow-up questions when needed?

Suttung Digital: Yes.

5.2 Summary

The interview reveals that there is no structured testing being done apart from internal testing of new features. This type of testing only reveals if the feature works as intended or if it has introduced new bugs. This means that Suttung Digital would benefit from having a framework that could help them find how well their game works from a potential customer's perspective.

The results produced during the user test provided beneficial data for several aspects of their game. Suttung Digital had already implemented several changes and features based on the recommendations from the test. As an example, they moved the player character higher up on the screen so that it was no longer obscured by the finger when using touch controller. They had also made the decision to ignore recommendations that went against their creative vision for the game. An example of this was the recommendation that the larger harmful objects in the games, such as the meteor, should do more damage. Overall, Suttung Digital's feedback was that the testing had been very beneficial.

Implementing the recommendations would likely take a couple of months in total. Some changes, like having text saying "click to start," would take very little time to implement. The larger things, like having more hazards in the game to create variation, would likely be done over a period of several months. This would be done alongside other development schedules for the game.

When asked if they could see themselves using the same methods that I had used, Suttung Digital felt that they should be able to use them. The biggest obsta-

cle mentioned would be the time consumption and the resources that need to be allocated for this. Recruiting test participants could also be a challenge, but they were confident that it would be doable. They use SCRUM with weekly sprints when developing their games. Testing on users after each sprint was not something they would envision doing, but they could see themselves doing testing after implementing new major features.

The heuristic rules did not seem too complicated, and it was mentioned that some of the rules listed are things that they are consciously aware of while designing the game. In their opinion, it would be good to have a list which provided them with a more structured way of evaluating their work.

It was believed that they would have been able to analyze the test results in the same way that was presented to them. However, it was stated that they would likely require a few attempts before they would be able to conduct the tests in a consistent manner so that the results would be reliable.

5.3 Outcome

The interview with Suttung Digital shows that the methods utilized during evaluation and user testing produce valuable data for small game developer studios. The methods are usable and understandable. It shows the need for these methods as the user testing that had been done by the developers themselves were sporadic and unstructured.

6 Discussion

6.1 Problem statement

As stated in Chapter 1, the goal of this thesis is to identify and test methods that would be usable by small game developer studios. This means that the methods have to be easy to use, require a low level of competence in usability testing, be easy to utilize, and not require a large sample size. The results produced in this thesis show that these requirements have been met by the methods that have been used for testing.

The literary review revealed that attempts have been made to create methods for testing games. However, these methods fail to comply with the needs of small game developer studios. They either require large sample size, specialized test facilities or competence in usability testing procedures.

6.1.1 Ease of use

The earlier studies in the field of playability testing did not look into how easy the methods are to use. This is in all likelihood because they were made for personnel with competence within the field. Earlier developed methods are outlined in Section 2.7 and from the overview of these methods it is clear that they have not been designed to be easy to use for someone with limited competence in using user-centered methods.

The experience gained from using the methods indicates that they are easy to conduct. No specialized equipment apart from a device for the game is required. The most complicated part of this was to write good notes. Recording sound or using screen capture software reduces this issue significantly.

The interview with Suttung Digital gave insight into how they perceived the tests. They concluded that they should be able to conduct these tests, but some experience would likely be required before the tests were consistent enough for reliable data gathering. This is to be expected when trying to utilize new methods in development.

6.1.2 Competence requirement

The impression from using the methods is that they are easy to learn, even for someone not familiar with conducting tests. The most challenging aspect is how to interpret test results. Suttung Digital did claim that they would likely come to the same conclusions as those presented to them from the *Fancy Diving* test, but there is no guarantee for this.

6.1.3 Small sample sizes

Small sample sizes did not hinder the data collection in any significant way. This corresponds well with Krug (2010) and Lewis (1990)'s view on small sample sizes, as shown in Section 2.9. The amount of data collected during the tests relied more on the complexity of the game rather than how many participated in the tests. The data collected from testing *Alex and the Search for the Elemental Stone* produced significantly more data than any of the other test, even though it had the smallest sample size. It only used three participants, whereas the other tests used five participants.

An important point made in Section 2.11 is to stop when you have found enough issues that would normally take the same amount of time to fix as there is time between testing. This limits the amount of test participants required.

6.2 Differences between the methods

There were noticeable differences between the methods when compared. A significantly higher amount of issues were found during the playtest of *Alex and the Search for the Elemental Stone*. Almost twice the amount when compared to the heuristic evaluation, 23 to 12. The type of game and its complexity may account for this. This is a larger game than the others that were tested, and there were less repeatable tasks. It also has a basic storyline that gives a clear indication of a starting and finishing point.

6.3 Results

The user-entered methods that were selected for use as part of the study were selected based on what had been done in earlier research and on criteria outlined in Section 3.1. This did show to work fairly well as they all produced relevant data when utilized. The small sample sizes used did not prevent the generation of significant amounts of data.

During the interview with Suttung Digital, it was revealed that the methods were something they could use in their game development cycles. The amount of issues found was significant enough for them to need several months to fully implement the suggested changes and features.

The amount of data generated when using the user-centered methods and the feedback from Suttung Digital show that the methods selected worked as intended.

6.4 Maintaining artistic integrity

The issue raised in Section 2.14 was how the methods used in this study could affect the game designer's artistic vision. In the study mentioned in the literature review, the designers managed to use the data from the user testing to get the game closer to their artistic vision.

Game testing tracks the user's usability issues, understanding, and perception

of the game. The game design intent is not directly tested, but the information gathered can reveal if the design has been successful.

In the interview with Suttung Digital, it is made clear that they decided against making changes that would have gone against what they want for their game. The test data recommended that specific objects in the game should do more damage as they looked more imposing. Suttung Digital made the decision to not implement this as it was in contrast to what they intended.

The feedback from Suttung Digital shows that game developers can ignore the recommendations from user tests. For this to be the case, the developers must be sure and confident in their design.

Even though Suttung Digital managed to utilize the test data generated in a way that maintained their artistic integrity, this may not be the case for all developer studios. Game design does not have one clear solution to an issue that appears during game testing, and less experienced developers may not know how to best deal with the feedback from the game testing.

6.5 Limitations of the study

6.5.1 Demographic

The games utilized in this study have not had a specified demographic. Based on this, a comparative study using a games specified demographic has not been conducted. It is, therefore, unknown if other results could have been obtained if more specific user groups had been used. Earlier studies have done the same thing. [Medlock et al. \(2002\)](#) specifically used people who had not played any games in the genre of the game they were testing. There are arguments made that say that users unfamiliar with a system are more likely to discover certain issues. Users familiar with a system may work around the issue. If this indeed holds true, then testing on a specific demographic may prove detrimental to the test results.

6.5.2 Development cycle

A full development cycle has not been observed. This is simply not possible due to the limited amount of time allocated for the thesis. This limits the knowledge of how effective the methods are during all stages of production and their viability for use over the full course of development.

6.5.3 Sample sizes and game complexity

The methods used may not be as effective for all games. Certain games may need more test participants before producing enough data for analysis. Game developers with little experience in testing may not be able to identify this problem, and can potentially end up making suggestions for change based on poor data. This may end up hurting the quality of the game.

There is also a potential issue with games that are more complex than those tested in this study, like strategy games. These games may require another ap-

proach were separate parts are tested individually in order not to get flooded with test results.

The game tests did reveal that less complex games may end up generating less valuable data than more complex games do. This becomes apparent when comparing the results from *Fancy Diving* with the results from *CityCop: Identify*. The approach for both games was similar, yet the amount of findings and recommendations was larger for *Fancy Diving*.

6.5.4 Full user experience

The games tested were pre-installed prototypes that required no installation on the part of the test participant. The full user experience was therefore not observed. For mobile and console games the installation process is the same for all installations, but this may not be the case for computer games. Certain services, like Steam, have unified the installation process for the games sold through their portal. There are still games sold as stand-alone products, both hard copies and through digital download. The installation process may have quirks that the methods in this study do not account for.

6.5.5 Platform differences

This study observed users that have played games on mobile phones, PC, and through a web-browser. Games on consoles have not been observed. Although there were no differences of note between how to observe the games used in this study, this may not be the case when playing on consoles. The approach differs somewhat, as the display for the game is usually further away than on the other gaming platforms. It may be harder to observe a test participant under these conditions.

[Virtual Reality](#) systems might suffer from the same issues when observing the player. It is possible to track what the player is seeing by using an external screen, but the sensory impressions are likely to be quite different.

6.5.6 Analyzing the results

Even though the tests show that data can be fairly easily gathered, it is not known if the results are interpretable to a game developer with little to no experience in interpreting these kinds of test results. It may hurt the design of the game if the test results are incorrectly interpreted.

6.5.7 Severity rating

The methods proposed do not employ any way of rating the severity of issues encountered, like the rating system outlined in Section 2.16. This is left to the game developer. They have to prioritize the issues themselves. This prioritization can potentially be based on what resources are available and not the severity of the issue. This could be detrimental to the final product.

In the interview with Suttung Digital, they stated that they have a discussion on what issues to prioritize.

As a full development cycle has not been observed in this study, there are no data available to substantiate any theory on this.

6.5.8 Occurrence rate

With samples as small as those used in the methods tested, there is the potential that issues with a low occurrence rate will not be discovered. How this will affect the final product is not known. There is reason to believe that this will not be as potentially harmful as it could be in traditional software. Bug testing and testing of general functionality can be done internally to some extent, at least for linear games and games that are limited in scope. Since games only produce intrinsic value, it is most important to utilize the tests to try to make the game as fun as possible.

6.6 Summary

The topics discussed in this Chapter show that the methods that have been utilized in this thesis are viable for their intended use. The methods are easy to use, they require little preparation and they are usable on a small sample size. The case study interview with Suttung Digital confirmed that they benefited from the use of the methods and that they would be able to use the methods themselves.

7 Conclusion

The goal of this thesis was to look at how user-centered design methods can be adapted for use by small game development studios. The literary review in Chapter 2 reveals that although work has been done in the field of game testing, there has not been a focus on developing methods for small game developer studios.

Three methods were selected as a result of the criteria outlined in Chapter 3. The methods needed to be usable with little preparation. The methods had to be usable on a small sample size. Finally, the methods had to be easy to conduct. The selection of methods was a heuristic evaluation, an interview, and an observation-style study called playtesting.

There was some variation in how well they performed. This variation was contributed to the differing complexity of the games tested. Of the four games tested in this study, two had a fair amount of complexity whereas the other two were more conceptual in nature. Chapter 4 reveals that the two complex games gathered larger amounts of data than the two conceptual games.

When the members of Suttung Digital were interviewed some weeks after they had gotten the test data for their game, *Fancy Diving*, it was clear that they had found the game test to be beneficial. They also stated that the methods used were something that they would like to use themselves.

The results in this thesis show that the methods used are effective at finding issues in games. There is variation as to how effective they are, but all methods did produce results with all of the games tested.

7.1 Suggestion for further research

There are several topics that could be further explored within the field of game testing. Many of the limitations of this thesis is outlined in Section 6.5. These limitations could be explored in future works.

The largest and what is likely to be the most beneficial area of study, is to observe a full development cycle of a game to measure the viability of the methods used during all the stages of development. This will provide insight into how much the studio relies on the methods for finding issues and how capable they are of conducting the tests and their ability to interpret the results. A study of this size is likely able to cover most of the limitations outlined in Section 6.5.

Bibliography

- Barwood, H. & Falstein, N. (2006), 'The 400 project'.
URL: <http://www.finitearts.com/Pages/400page.html>
- Beyer, H. (2010), *User-centered Agile Methods*, Synthesis lectures on human-centered informatics, Morgan & Claypool.
URL: <http://books.google.no/books?id=CoYxZ7mWapUC>
- Clanton, C. (1998), An interpreted demonstration of computer game design, in 'CHI 98 Conference Summary on Human Factors in Computing Systems', CHI '98, ACM, New York, NY, USA, pp. 1–2.
URL: <http://doi.acm.org/10.1145/286498.286499>
- Davis, J. P., Steury, K. & Pagulayan, R. (2005), 'A survey method for assessing perceptions of a game: The consumer playtest in game design', *Game Studies: The international journal of computer game research* 5(1).
- Desurvire, H., Caplan, M. & Toth, J. A. (2004), 'Using heuristics to evaluate the playability of games', *CHI '04 Extended Abstracts on Human Factors in Computing Systems* pp. 1509–1512.
URL: <http://dl.acm.org/citation.cfm?id=986102>
- Desurvire, H. & Wiberg, C. (2009), 'Game usability heuristics (play) for evaluating and designing better games: The next iteration', *Online Communities and Social Computing* 5621, 557–566.
- Digital Anvil (2003), 'Brute force', DVD-ROM.
- Entertainment Software Association (ESA) (2015), '2015 sales, demographic and usage data - essential facts about the video game industry', Web.
URL: <http://www.theesa.com/wp-content/uploads/2015/04/ESA-Essential-Facts-2015.pdf>
- FASA Interactive (2007), 'Shadowrun', DVD-ROM.
- Federoff, M. A. (2002), Heuristics and usability guidelines for the creation and evaluation of fun in games, Master's thesis, Indiana University.
- Fulton, B. & Medlock, M. (2003), 'Beyond focus groups: Getting more useful feedback from consumers'.
URL: www.itu.dk/ljc/speciale/tekster/Beyond_focus_groups.doc

- Kim, J. H., Gunn, D. V., Schuh, E., Phillips, B., Pagulayan, R. J. & Wixon, D. (2008), 'Tracking real-time user experience (true): a comprehensive instrumentation solution for complex systems', *CHI '08 Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* pp. 443–452.
URL: <http://dl.acm.org/citation.cfm?id=1357126>
- Krug, S. (2010), *Rocket Surgery Made Easy: The Do-it-yourself Guide to Finding and Fixing Usability Problems*, Voices that matter, New Riders.
URL: <https://books.google.no/books?id=VzbimAEACAAJ>
- Lazar, J., Feng, J. & Hochheiser, H. (2010), *Research Methods in Human-Computer Interaction*, John Wiley & Sons.
URL: https://books.google.no/books?id=H_r6prUFpc4C
- Lewis, J. R. (1990), 'Sample sizes for observational usability studies: Tables based on the binomial probability formula', *Tech Report 54.571* .
- Lewis, J. R. (1991), 'Legitimate use of small samples in usability studies: Three examples', *Tech Report 54.594* .
- Lieng, A., Lunde, H. E., Martinsen, M. A., Larsen, R. & Forsberg, J. B. (2013), 'Alex and the Search for the Elemental Stone'.
- LucasArts (1990), 'The Secret of Monkey Island'.
- LucasArts (1993), 'Day of the Tentacle'.
- Malone, T. W. (1981), 'Toward a theory of intrinsically motivating instruction', *Cognitive Science* 4 pp. 333–369.
- Malone, T. W. (1982), Heuristics for designing enjoyable user interfaces: Lessons from computer games, in 'Proceedings of the 1982 Conference on Human Factors in Computing Systems', CHI '82, ACM, New York, NY, USA, pp. 63–68.
URL: <http://doi.acm.org/10.1145/800049.801756>
- Medlock, M. C., Wixon, D., Terrano, M., Romero, R. & Fulton, B. (2002), 'Using the RITE Method to Improve Products: a Definition and a Case Study', *Usability Professionals Association* .
- Microsoft (1999), 'Age of Empire II'.
- Microsoft Game Studios (2001), 'Halo: Combat evolved', DVD-ROM.
- Microsoft Game Studios (2004), 'Halo 2', DVD-ROM.

- Nielsen, J. (1995), 'Severity ratings for usability problems'.
URL: <https://www.nngroup.com/articles/how-to-rate-the-severity-of-usability-problems/>
- Pagulayan, R. J., Keeker, K., Fuller, T., Wixon, D., Romero, R. L. & Gunn, D. V. (2007), 'User-centered design in games', *Human-Computer Interaction Handbook*.
- Pinelle, D., Wong, N. & Stach, T. (2008), Heuristic evaluation for games: Usability principles for video game design, in 'Proceedings of the SIGCHI Conference on Human Factors in Computing Systems', CHI '08, ACM, New York, NY, USA, pp. 1453–1462.
URL: <http://doi.acm.org/10.1145/1357054.1357282>
- Rubin, J., Chisnell, D. & Spool, J. (2011), *Handbook of Usability Testing: How to Plan, Design, and Conduct Effective Tests*, Wiley.
URL: https://books.google.no/books?id=l_e1MmVzMb0C
- Sánchez, J. L. G., Vela, F. L. G., Simarro, F. M. & Padilla-Zea, N. (2012), 'Playability: analysing user experience in video games', *Behaviour & Information Technology* **31**(10), 1033–1054.
URL: <http://dx.doi.org/10.1080/0144929X.2012.710648>
- Shippy, D. & Phipps, M. (2010), *The Race For A New Game Machine: Creating The Chips Inside The Xbox 360 And The Playstation 3*, Kensington.
- URL:** <https://books.google.no/books?id=FYcym0LT3CYC>
- Shneiderman, B. (1986), *Designing the User Interface: Strategies for Effective Human-computer Interaction*, Addison-Wesley Longman Publishing Co., Inc., Boston, MA, USA.
- Sommerville, I. (2011), *Software Engineering*, International Computer Science Series, Pearson.
URL: <http://books.google.no/books?id=l0egcQAACAAJ>

Glossary

accessibility How accessible a product is in order to accomplish a goal. [14](#)

effectiveness How well the product behaves compared to the user's expectancy of how it should work. [14](#)

efficiency How quick a user's goals can be accomplished accurately. [14](#)

learnability How easy it is to learn how to use the product, related to effectiveness. [14](#)

playability A measure of either the ease by which a video game may be played, or of the overall quality of its gameplay. [2](#), [14](#)

Real Time Strategy A genre of strategic video games where the player moves units in real time as opposed to turn-based movement. [11](#)

satisfaction The user's perceptions, feelings and opinion of the product. [14](#)

usefulness The degree to which a product enables users to achieve their goals and their willingness to use the product. [14](#)

Virtual Reality is a computer technology that replicates an environment, real or imagined, and simulates a user's physical presence and environment to allow for user interaction. [45](#)

A Alex and the Search for the Elemental Stone - Test Results

A.1 Heuristic Evaluation

A.1.1 Heuristics

The heuristics used are derived from the work of [Desurvire et al. \(2004\)](#) and [Desurvire & Wiberg \(2009\)](#). Heuristics that are not applicable to the selected game have been excluded from the heuristic evaluation. The game is first played and then the heuristics are applied. How well the game conforms to a selected heuristic is identified. Suggestions are then written that explains what is necessary in order to make the game conform to said heuristic.

The heuristics are split into four different categories; Gameplay, Game Story, Mechanics and Usability.

Gameplay

Table 6: Heuristic rules for Gameplay

	Heuristic and Description
1	Player's fatigue is minimized by varying activities and pacing during game play.
2	Provide clear goals, present overriding goal early as well as short-term goals throughout play.
3	The game is enjoyable to replay.
4	Player should not experience being penalized repetitively for the same failure.
5	Players should perceive a sense of control and impact onto the game world. The game world reacts to the player and remembers their passage through it. Changes the player makes in the game world are persistent and noticeable if they back-track to where they've been before.
6	The first player action is painfully obvious and should result in immediate positive feedback.
7	Pace the game to apply pressure but not frustrate the player. Vary the difficulty level so that the player has greater challenge as they develop mastery. Easy to learn, hard to master.
8	Challenges are positive game experiences, rather than a negative experience (results in their wanting to play more, rather than quitting).

Game Story

Table 7: Heuristic rules for Game Story

	Heuristic and Description
1	Player understands the story line as a single consistent vision.
2	Player is interested in the story line. The story experience relates to their real life and grabs their interest.
3	Player experiences fairness of outcomes.
4	The game transports the player into a level of personal involvement emotionally (e.g., scare, threat, thrill, reward, punishment) and viscerally (e.g., sounds of environment).
5	Player is interested in the characters because (1) they are like me; (2) they are interesting to me, (3) the characters develop as action occurs.

Mechanics

Table 8: Heuristic rules for Mechanics

	Heuristic and Description
1	A player should always be able to identify their score/status and goal in the game.
2	Mechanics/controller actions have consistently mapped and learnable responses.
3	Controls should be intuitive, and mapped in a natural way; they should be customizable and default to industry standard settings.
4	Player should be given controls that are basic enough to learn quickly yet expandable for advanced options.

Usability

Table 9: Heuristic rules for Usability

	Heuristic and Description
1	Provide immediate feedback for user actions.
2	The Player can easily turn the game off and on, and be able to save games in different states.

3	The Player experiences the user interface as consistent (in control, color, typography, and dialog design) but the game play is varied.
4	The Player should experience the menu as a part of the game.
5	Upon initially turning the game on the Player has enough information to get started to play.
6	Players should be given context sensitive help while playing so that they do not get stuck or have to rely on a manual.
7	Players do not need to use a manual to play game.
8	The interface should be as non-intrusive to the Player as possible.
9	Make the menu layers well-organized and minimalist to the extent the menu options are intuitive.
10	Art should be recognizable to player, and speak to its function.

A.1.2 Evaluation

Gameplay

Table 10: Heuristic evaluation of Gameplay

	Evaluation result
1	Sections of the game are finished by completing a “mini-game” which differs substantially from the rest of the gameplay.
2	The goal of the different worlds is not specified in any manner. The “mini-games” have an instruction before they start that explains what to do.
3	The game is very linear. As such, there is little reason to play it again.
4	The main game sections are inherently built in a way that does not penalize mistakes. Some of the “mini-games” can be penalizing and frustrating.
5	What the player has done is clearly visible. When a section of the game has been completed it is graphically displayed.
6	Figuring out how to move the character in the game is easy.
7	The main gameplay sections have no distinct variation in difficulty. There are large differences in the difficulty of the different “mini-games”. There is no pacing of note.
8	Some challenges can be very frustrating as there are no indications as to what the player should do.

Game Story

Table 11: Heuristic evaluation of Game Story

	Evaluation result
1	There is an introductory movie showing the background of the player character. Apart from this, there is no story of note.
2	Story feels uninteresting and unengaging.
3	Outcome seems fair.
4	No
5	The only persistent character in the game is the player character. He is unlikely to be like the person playing the game. He has no personal development over the course of the game. He feels uninteresting. However, he is drawn in such a way that he might be appealing regardless of these shortcomings.

Mechanics

Table 12: Heuristic evaluation of Mechanics

	Evaluation result
1	There is no score in the game. Progression can be seen on the world map, but this is only visible once a section of the game has been completed. The player is able to see what items have been picked up.
2	The main portions of the game have consistent control methods that feel natural. There are some minor variations in the “mini-games”, but the are not of any hindrance and they are well explained.
3	Controls are natural and they are mapped in a way that conforms to industry standards. It is not possible to remap controls.
4	The controls are ease and quick to learn. There is no need for expanding controls for more advanced features in this game.

Usability

Table 13: Heuristic evaluation of Usability

	Evaluation result
1	Picking up and using items are immediately shown.

2	The game is easy to close. The player can select which portions of the game that has been completed earlier. This must be done from the main menu.
3	The interface does not change over the course of the game. Gameplay is varied.
4	There is no in-game menu. The main menu on the start screen feels as a part of the game.
5	There is some chance that novice gamers will not understand the gameplay. There are instructions in the menu that is displayed on the start screen of the game.
6	Very little contextual information is given.
7	It is unlikely that a manual is needed in order to play the game given the simplistic nature of the game.
8	The interface is non-intrusive.
9	The start menu is well organized and minimalistic.
10	Art in the different sections of the game stands out when needed. Their purpose is not always clear.

A.1.3 Recommendations

The following are recommendations based on the findings of the heuristic evaluation of “Alex and the Search for the Elemental Stone” and general impressions of the game during this evaluation.

Gameplay

What to do in the different sections of the game could be improved upon by having more contextual information that gives clues as to what the player should do next. As an example; the elf in the start area of the Earth level should display a text box that explains that he has lost his flower and that he wants the player to find it for him. Generally, NPCs should give some sort of feedback when the player tries to interact with them.

The linearity of the game is inherent in the genre and there is not much that can be done about this. Having difficulty levels on the “mini-games” could help to alleviate some of the frustrations that are likely to occur. This is especially the case with the end-game of the Ice level, which is very punishing even for the smallest mistake.

There is no pacing of note in the game, but in adventure games of this type it is not uncommon.

Game Story

The story of the game is non-existent for most of the game, apart from the game’s introductory movie. Having some sort of introduction movie and ending to every zone could help to improve on this. The end level should especially have a movie that is being played when completing the game.

More characters should be created, or the existing ones should be fleshed out more, and they should have repeat appearances throughout the game. A

potential love interest for the player character could help to provide a more engaging narrative and to make the player character more likeable. The wizard should also make an appearance outside of the introductory movie.

Mechanics

Older adventure games often had a score system where it is possible to see how many points the player has earned of a total. This can help the player to keep track of the progression, but this is strictly speaking not a necessity. The controls are so simple that there is really no need to offer remapping of the control scheme.

Gameplay

There is no way to return to the main menu once the player has started the game. This is due to the fact that there is no in-game menu available. This should be implemented and an option to exit the current game should be made available. Save and load options can also be implemented here.

The interface is generally OK, but there are some minor bugs when interaction with objects in the inventory. In some cases, the player character moves to the position of an item in the inventory when the player clicks on it. Another minor bug is items in the inventory not moving as they should when they are used.

The gameplay instructions in the main menu helps players to understand what to do, but there should also be a tutorial early in the game showing how this works in practice.

General Notes

The game feels like it is in an unfinished state at the moment. There are elements that need polishing. This is especially true for the movement system as the player character sometimes get stuck in certain areas.

A.2 Play session

Observe the test participant and make notes of issues and bugs that show up during gameplay. Ask the test participants to talk out loud when a problem is encountered or there are other issues that they do not understand.

Tell the test participants to play until they complete the game or they do not feel like playing the game anymore, for whatever reason. If they quit before the game is completed ask them what caused them to end the game.

Questions the player has about the game must not be answered until the play session is over. Make this clear before the game session begins.

The notes taken are written into longer narratives after completion of the play sessions. This is done in order to provide context for anyone not participating in observing the play sessions.

A.2.1 Test Participant 1

Male, 25, Experienced Gamer.

Understood the world map right away. Started with Air level.

Air Level

First area was not clear on what to do. Tried clicking on the umbrella and eagle, with no effect. Went to the second area and picked up the three clouds. Went back to the first area. Saw the three clouds on the umbrella and used the clouds from the inventory on it in order to complete the air level. The mini-game was understandable, but didn't understand the reasoning behind shooting the two people at the bottom of the screen.

Selected Earth level from the world map. Noticed that the air level was graphically changed and understood that it indicated that it was completed.

Earth Level

The first area of the earth level did not give any indication as to what was needed to do. Tried clicking the elf and the door on the left side on the screen several times before moving on to another area. During this, the player got stuck on the trees several times. The player then went to the left area and tried clicking on the rabbit. He figured that certain items were needed in order to be able to progress. The next area visited was the middle area. Here the player tried clicking on the flower and the carrot without result. He then moved on to the right area. During the time he went to different areas, he got stuck on trees in the first area several times, causing frustration.

In the right area the player immediately picks up the bucket and the spade. After a few failed attempts at using the bucket on the well, he went back to the middle area and used the shovel on the carrot and then water on the flower. He then went to the left area and used the carrot on the rabbit and picked up the weeds. He then tried to use the weeds on the elf in the first area and then went to the right area and used the weeds on the goat. There was a brief period where he did not understand what to do, but then used the shovel on the dirt under the rabbit's trunk. He then used the flower on the elf, completing the level.

During the play through of this level the player tried to close the animation window by clicking outside the animation window. Figured out how to close it immediately after doing this once.

Next level selected was the Water level.

Water Level

Started by going to the north area from the start area. Here he tried clicking on the fisher and the bottle which gives no effect. He then went to the right area and clicked on the snowman. Was confused by the information the snowman gave.

He then went to the north area and gave the worms from the snowman to the fisherman. He then went to the south area and clicked on the ice block. Liked the animation played when doing this. Went to the right area and used the bottle on the snowman.

The player now went to the start area and tried clicking on the seal. Found the graphics confusing as there was no interaction with the seal.

In the south area the player now combined the tongue with the fishing rod and got the last part needed in order to complete the lever. Enjoyed the thought process behind this puzzle. Went to the start area and completed the level.

The mini-game was somewhat frustrating and there were several occurrences of vocal frustration with the game. Once completed, the next level selected was Fire.

Fire Level

The text displayed when clicking the NPC in the first area was hard to read. Picked up the fire shard when the NPC disappears. Went to the second area and picked up the rope. In the third area he picked up the stick and in the fourth area he picked up the hook. When trying to combine the items, he encountered a bug where the items in the inventory started to overlap each other. Solved this by clicking on the items and trying to move them. This fixed the issue, even though this is not an intended function.

In the third area he used the rope with the hook in order to get the fire shard. And in the second area the stick was used to get the last shard. He then went to the fourth area and used the fire emblem on the door's lock. Wondered if the designers got lazy at this point, as he thought the level was too easy. Thought that the mini-game was too easy as well. Complained about bad "A.I.".

There was a short confusion on what to do on the world map when the four main levels are marked as completed. Saw the mountain was not marked in and clicked on it, starting the end level.

End Level

Complains about imprecise jump mechanics, especially that the character sometimes hovers a bit over the ground. Did not like that the character does not stop immediately after releasing the directional buttons. Experienced instances where it was impossible to avoid the fire ball.

When completing the level, the player reacted to the fact that the game does not offer an option to return to the main menu.

Other comments

Had to switch between keyboard and mouse when restarting the mini-games.

A.2.2 Test Participant 2

Female, 23, moderately experienced gamer.

Started by reading the instructions in the start screen menu. Tried to click on the screen during the introduction. Clicks continue when the introduction has finished playing.

Water Level

Tries to go to the northern area, but clicks to low to trigger the transition. Unsure of what is happening. Tries to click again, this time higher up, which triggers the transition to the new area. Tries to click on elements in the area in order to see what happens. Moves on to the southern area when nothing happens. Clicks on the Ice-block, and tries to click on the icicle. Moves on to the eastern area when nothing happens. Clicks on the snowman and then tries to use the tongue on the snowman. During several tries of this, the inventory items starts to move around a lot, causing some confusion. Moves on to the northern area and tries randomly to use items on different objects in the area without success. Eventually uses the worm on the fisherman. After this the flask is used on the snowman. At this point the inventory system bugs out and the game needs to be restarted.

After the game is restarted, the puzzles are solved up until the same state as before the bug. After trying many different options, a hint is given as to where the fishing rod with the attached tongue can be used.

Mini-game is completed fairly quickly after a few tries.

Air Level

Walks to the right screen and picks up the clouds. Goes back and tries to click on the eagle. Uses the clouds on the umbrella. Mini-game is completed quickly

Earth Level

Starts walking to elf in order to try to interact. Walks to the left of the elf and gets stuck between the door and the elf. Manages to get loose by clicking randomly. In the left and right area the player tries to click on the objects. Goes to the right area and picks up the spade and bucket. Tries to use the bucket on the well. Walks back to the start area and gets stuck on the trees. Figures quickly out how to get loose. Uses shovel on carrot and carrot on rabbit. Picks up weeds. Walks back to the start area and expresses uncertainty as to what the elf does. Tries to use bucket on the flower, but misses the graphic when clicking and nothing happens. Tries to click random items in the inventory, including the flower petals, at which point the items moves around in such a way that they cover each other. The spade is covered by a flower petal. Uses the spade on the dirt in the left area. After a while with not figuring out what to do, the player is told that the initial use of bucket on the flower was correct. Uses the flower on the elf. Mini-game is easily solved.

Fire Level

Must lean forward in order to read the NPC text. Picks up the fire shard and moves on to the next area. Picks up the rope, tries to pick up the fire shard. Moves on and picks up the stick and the hook. Combines rope and hook and uses it on the fire shard then uses the stick on the last fire shard. Uses the fire emblem on the door lock. Text before the mini-game is hard to read. Solves

mini-game quickly.

End Level

The instruction text is hard to read. Gets frustrated by uneven gameplay with fireballs that can't be avoided and imprecise controls. Quits the game after some tries.

Other comments

Double-clicks items in the inventory when trying to select items. Using some of the inventory items is difficult as the test participant does not always click directly on the graphics. Tries to use inventory items that are not clickable. Has to use mouse to restart the mini-games.

There was a big jump during the play session in the understanding of how the game worked. This caused the later levels to be solved quicker with less hassle.

A.2.3 Test Participant 3

Male, 18, Experienced Gamer

Started without reading instructions. Skips introduction before it is completed. Quickly figures out how to use the world map.

Water Level

Goes right from the start area. Clicks on snowman and gets the worm. Goes to the south area and clicks on ice block. After a little while the player figures out that he can use inventory items in order to interact with objects in the game. After some attempts at trying to use the inventory objects at random, he uses the worm on the fisherman and gets the fishing rod and bottle. Goes east to the snowman and uses the bottle.

In the start area he tries to use the random objects on the seal. He does the same thing in the southern area, where he tries using items on the icicle. After a while it becomes apparent that he is not aware of the option to combine items in the inventory. In order to make progress, he is told that this is a possibility. The combined tongue and fishing rod is used on the fishing hole beside the fisherman, and on the seal before moving to the southern area where it is used on the icicle. The icicles are then used in order to finish the level.

The mini-game is found to be quite hard and he states that he would quit the game after a certain amount of attempts to complete it. It is also stated that he enjoys that there are different versions of the labyrinth as it keeps it interesting.

Earth Level

Initially tries to talk to the elf in the first area. Goes to the left area first. Tries to interact with the rabbit. Moves on to the middle area and tries to interact with the flower and carrot. Gets stuck when exiting this area, but figures out how to get unstuck. Moves on to the right area and picks up the bucket and the spade. Tries to use the bucket on the well. In the middle area he uses the water bucket

on the flower and the spade on the carrot. Goes to the left area and uses the carrot on the rabbit. He then uses the spade on the dirt under the rabbit. Clicks on the weeds, but does not hit the graphic directly and it is not registered by the game. At this point he expresses that he finds the puzzles in this are more natural/logical than they were in the water level.

There is now some time where he is unable to figure out what to do. Gets told that it is possible to pick up the weeds. He does this and moves to the right area in order to use the weeds on the goat. Misses the graphic in the inventory a couple of times before he is able to use the item. He then gives the flower to the elf. The mini-game is then solved easily.

Air Level

Goes to the right area and picks up the three clouds. Moves back to the left area and tries to use the clouds on the eagle before using them on the umbrella. Solves the mini-game without any issues.

Fire Level

Solved the NPC questions easily. Picks up the shard and moves the next area where he picks up the rope. Continues moving right while picking up all the items. Reaches the last area and combines the hook with the rope. Goes back and uses the rope and hook on the shard on the mountain. Uses the stick on the fire shard in the lava. Uses the fire emblem on the door.

Easily wins the first game. Does not figure out how to start a second round of games. Tries to click randomly around the screen, but misses the game itself. After a while of trying to figure this out he is told how to start a new round. Easily solves the last two rounds.

End Level

Comments that he likes the game as it reminds him of similar games that he is familiar with, such as Super Mario Bros. Does not like the random fireballs, as they are sometimes unavoidable. Would like the option to duck in order to help to avoid the fireballs. Makes a remark that flowers he has been hit by does not come back.

Other comments

Had to use the mouse in order to restart some of the mini-games, thereby switching between keyboard and mouse. Used the continue button in the introduction because it was present.

A.2.4 Recommendations

This section contains suggestions for changes in the different levels and general suggestions for improving the game experience based on player observation.

Earth Level

Participants favored going to the left area first, where they could not figure out anything to do. Switching the left and right area will help players to find items that they can use quicker. The current right area contains the first items that the player needs to pick up in order to proceed.

All participants tried to use the water bucket on the well, this confusion could possibly be solved by either having an empty bucket that gets filled up by the well or by having a full bucket hanging in the rope of the well that the player needs to pick up.

Participants tried talking to the elf in the start area, giving him a text box that explains that he has lost his flower, and that he will open the door if the player retrieves it for him should be added. A bug which allowed the player to walk through the door and end up behind the elf was observed. It should be made sure that this is not possible. Getting stuck between trees when moving between areas was a common occurrence.

One participant noted that he felt that this area had a more natural flow to it than the other levels.

The mini-game is easily solved by all participants.

Air Level

This area was easily solved by all participants. There was some confusion with one participant, but this is in all likelihood due to the fact that this was the first level he selected.

This level is probably too short and straight-forward. More puzzles and complexity should be added in order to make it a more enjoyable and lasting experience.

The mini-game was solved fairly easy by all participant. One participant did note that he did not understand the reason for why he was shooting lightning at the characters on the bottom of the screen. A story line explaining this would be helpful.

Water Level

The seal in the start area was confusing as it was thought to be an interactive NPC. It should either be made clear that it is there for aesthetic purposes or it should be removed.

The puzzles in this area are to a large degree solved by random use of items. One participants noted that the text box displayed when clicking the snowman was confusing.

The mini-game can be very frustrating. The smallest mistake made results in having to restart. Consider lowering the difficulty level of this mini-game. The different layouts of the labyrinth was welcome.

Fire Level

The text displayed on the NPC in the starting area was hard to read for some of the participants. The font and color of this should be changed in order to make it easier to read.

This level is too straight forward. All participants went through the level without any problems at all. More puzzles and complexity should be added in this level. One participant said he thought this area felt lazy.

The mini-game itself was very easy. One participant complained about it being too easy as the "A.I" provided no challenge at all. Another participant was unsure as to how to start a new round in the mini-game as there is no button or any other hints as to how to do this. A button to the left of the mini-game itself with the text "Next round", or something similar, would be beneficial here.

End Level

The main issue with this level is the poor control mechanics. The character was hard to control precisely. At certain time the random fireballs appeared in a position which made it impossible to avoid them. One participant requested the ability to duck so that it would be possible to avoid these fireballs.

The main character sometime hovers over the ground instead of standing in the ground. The movement animation sometimes fails to display the correct animations.

General Recommendations

Problems using the inventory were observed several times. This is in part due to a bug where inventory items moved around and sometimes ended up on top of each other. In one instance this resulted in an inability to progress in the game, forcing a restart. This is a potentially game breaking issue and should be given a high priority.

At several instances during the gameplay session it was observed that items did not interact correctly when used on the correct objects in the game world. This is due to the graphics being "click-through" in the areas that are not occupied by graphics. As an example, in the Earth level, when using the water bucket on the flower in the middle area, one player used the correct item, but clicked in the open area between the hanging flower and the stem. This should be solved by making a larger area outside the specific graphics clickable.

All participants were a little lost in the first level they selected, this indicates that they learned much of how the game worked through the first level played. A small tutorial before the main game might help to familiarize players with how to play the game. The ability to combine items was not obvious to all participants.

The mouse must be used in mini-games that requires a restart in order for the participant to start a new attempt. This should be changed so that the keyboard can be used to restart the game. A text book should say something like "Press 'R' to restart".

Generally the text in the game is hard to read, with the exception of the main menu.

A.3 Post-interview questions based on heuristics

After the gameplay session, conduct an interview. The questions asked are based on the heuristics used in phase 1. The approach for the interview is a semi-open structure. The questions are asked as they are written, but follow-up questions are encouraged in order to expand upon the answers given.

A.3.1 Questions and answers

Questions are marked with a number corresponding to the heuristic rule within each of the four categories. The answers are numbered in a way that corresponds to which test participant provided the answer. Answers are reproduced in a succinct manner, not verbatim.

Gameplay

1. Was there enough variation in gameplay?
 1. Not mechanically, much of the game was the same. The mini-games helped to create more variation.
 2. There was enough variation.
 3. Yes, liked new levels and finding different items in each level.
2. Did you understand what you were supposed to do?
 1. Most of the time, but as a very experienced adventure game player this comes naturally. More textual aids might be good. Graphics helped to understand what to do in certain areas.
 2. Not at the beginning, understood better as progression was made in the game.
 3. No, not at the beginning of the game.
3. Would you like to play the game again?
 1. No, the game is too linear.
 2. Yes, if the graphics get improved.
 3. Yes, parts of the game, especially the end level.
4. Did you feel that you were punished for trying the same thing several times?
 1. Not particularly. The last level felt punishing.
 2. Yes, the labyrinth mini-game and the final platform game.
 3. No, some of it was hard, but not punishing.
5. Were your actions in the world noticeable to you?
 1. Yes.
 2. Only that you can see what you have collected of the items.

3. Yes, noticed changes when tasks were completed.
6. N/A
7. Was the difficulty level of the game frustrating, decent or good and were there too large jumps in difficulty?
 1. Not harder than other games. The mini-games were the most challenging part of the game.
 2. Yes. But it may be that there are limitations in the game rather than the difficulty level. The final platforming level has too large a jump in difficulty level.
 3. It was fine. It was hard until I got the feel for how the game works.
8. Did the challenges in the game make you want to play more?
 1. Yes, mainly to complete the game. Considers himself a “Completionist”.
 2. Yes, until the last part at least.
 3. Yes, got the desire to complete all the levels in the game.

Game Story

1. Did the storyline feel consistent?
 1. Was OK. No indication of how the world works.
 2. No, but it felt fine within each of the levels. There was a natural progression.
 3. Hard to say as I skipped the intro. Got a better feel for the story after watching it.
2. Was the storyline of interest to you?
 1. Not particularly.
 2. Not for people of my age, might be fine for younger people.
 3. Hard to say. I think that there should be a fixed order that the game needs to be completed in.
3. Did the outcome of the challenges satisfy you?
 1. Didn't get the point.
 2. As long as I was winning.
 3. Yes, feels like I've been through a journey.
4. Were the environments immersive/enjoyable?
 1. They were OK.
 2. Yes, got a feeling that the levels were seasonal.
 3. Got very immersed into the world. Liked the different settings.
5. Did you like the main character? What about the other characters?
 1. Yeah, could relate. NPCs were OK, the non-interactive NPCs can be

confusing (eagle, seal).

2. Yes, reminds me of "Espen Askeladd"¹. The fisherman was OK. Liked the wizard.
3. Feels a connection with the main character. The elf stood out compared to the other NPCs.

Mechanics

1. Did you feel that you had a good understanding of your status in the game?
 1. Only inventory gave any indication of change. Didn't know what had been completed.
 2. No, did not have a good overview. The arrows helped. Didn't know what was left to do.
 3. Yes, get the feeling that you want to help (might have misunderstood the question).
2. Was the control system easy to learn?
 1. Easy to understand.
 2. Yes.
 3. Yes, old and familiar controller scheme.
3. Covered in question 2.
4. Covered in question 2.

Usability

1. Was the feedback to your actions immediate enough?
 1. Problem was realizing what to do. Feedback was good. Some note of what elements in the game was interactive would have helped.
 2. Yes and no. Bad feedback when combining items.
 3. Yes, enjoyed the cut-scenes that confirmed actions.
2. Did the user interface (different wording) change in ways you did not understand?
 1. No.
 2. No.
 3. Yes, had some difficulty understanding that items were not drag-and-drop.
3. Did the start menu feel consistent with the rest of the game?
 1. Some of the elements were too graphically different. The text and close buttons were not the same in all of the different levels.
 2. Yes.
 3. Yes, gives a feeling of what the game is about.

¹<https://en.wikipedia.org/wiki/Askeladden>

4. Did you have enough information to start playing the game right away?
 1. Yes.
 2. Reading the instructions helped understanding the game.
 3. Yes, until the first challenge in the game. Controllers were easy to understand.
5. Did you feel like you could have used a manual at any stage of the game?
 1. Textual aids would have been helpful.
 2. Yes, in the ice world. This was early in the game, had a better understanding later in the game.
 3. No, but a text box or something that explains what you have to do would have been helpful.
6. Did the graphical user interface(wording) interfere with your gameplay?
 1. No.
 2. It was not interfering.
 3. In the ice-level, yes.
7. Were the interface options easy to understand?
 1. Some minor bugs interfered, design was good.
 2. Yes.
 3. Understood them quickly.
8. Were the graphics confusing or did they help you to understand what they were meant to do in the game?
 1. Arrows on the ground indicating where the player could go was helpful.
 2. The arrows was very helpful. The items that needed to be picked up were not distinctive enough.
 3. They helped, saw quickly what you could use.

A.3.2 Recommendations

These recommendations is based on the overall impressions of the feedback within each category of questions.

Gameplay

There is indication that the participants did not initially understand how the game worked. A tutorial at the beginning of the game might help to alleviate this problem. As would more textual aids in the game.

Game Story

The overarching storyline did not give the participants a good feeling for what is happening in the game. It was not of much interest. Animations or at least text that describes what is happening between each level could help to improve on

this.

All the participants enjoyed the settings in the different levels. The main character was generally well liked, but the NPCs did not leave much on an impression.

Mechanics

Their current status in the game world not apparent to the participants. More information and indications as to what is left to do should be implemented.

The controller scheme was easy to understand.

Usability

It is not apparent what objects in the world is interactive. A function that indicates this could be an outline of an object that is interactive when moving the mouse pointer over it.

One participant noted that the graphical presentation of some elements in the game differed too much. Some tightening of the graphical elements in order to provide a more graphic consistency should be considered.

One participant indicated that a manual could have been useful at the early stages of the game, whereas the others would have liked textual aids. NPCs should have text boxes that gives some indication as to what the user is supposed to do.

With the exception of some bugs experienced in the user interface, it was easily understandable. Graphical aids in the game, specifically the arrows indicating where the player character could go, was helpful.

B Fancy Diving - Test Results

B.1 Heuristic Evaluation

The heuristics used in the pilot studies were split into categories. This categorical division will be reused with some modifications. The story category will be removed as the game has no story of note.

The heuristic evaluation was performed on "Fancy Diving" version 0.1 #17.

B.1.1 Heuristics

Gameplay

Table 14: Heuristic rules for Gameplay

	Heuristic and Description
1	There is enough variation in gameplay to keep the player interested.
2	The gameplay elements are consistent in shape and function.
3	The difficulty of the game is well paced.
4	It is clear to the player what should be done during the different stages of the game.
5	Customization of the character through visual changes or power-ups should be available in order to help immerse the player.
6	Challenges in the game are enjoyable and keeps the player wanting to play more rather than becoming frustrated.

Mechanics

Table 15: Heuristic rules for Mechanics

	Heuristic and Description
1	Controllers are consistently mapped.
2	The player's status in the game is always identifiable.
3	The game controllers follows industry standards and conventions, with the exception of the third stage.

Usability

Table 16: Heuristic rules for Usability

	Heuristic and Description
1	The art in the game should give a clear indication of what they are and their purpose in the game.
2	The player does not require a manual or an in-game tutorial in order to understand how to play the game.
3	The user interface is consistent.
4	The main menu of the game feels like a part of the game.
5	There is enough information available in order to start playing the game as soon as it has loaded.
6	Information on how to play the game is available during gameplay.
7	Sounds and visual effects gives gameplay feedback to the player.
8	Interface elements do not interfere with gameplay.

B.1.2 Evaluation

Gameplay

Table 17: Heuristic evaluation of Gameplay

	Evaluation result
1	The gameplay itself has little variation in gameplay. The first and second phase of the game are fairly similar, with the exception of gathering coins in the second phase. The third phase differs substantially from the the other two phases.
2	Gameplay elements are consistent.
3	The difficulty level increases as the player progresses in the first phase. There is no feel of an increased difficulty level in the second and third phase. The only difference experienced is that they last longer.
4	The first and second phase are quite easy to understand. The third phase suffers from having bad interface feedback on how the user is supposed to perform the different tricks.
5	There is no option to customize anything in the game.
6	The challenges, with the exception of the third stage are not frustrating. However, they do not encourage the player to continue playing either.

Mechanics

Table 18: Heuristic evaluation of Mechanics

	Evaluation result
1	The first and second phase of the game have identical controls. The third phase changes the controls. This is not a problem however, as it is quite obvious to the player that this is a different game mode.
2	Status is clearly and consistently displayed.
3	The controllers follow industry standards, with the exception of the third stage.

Usability

Table 19: Heuristic evaluation of Usability

	Evaluation result
1	The clouds that appears as dangerous objects later in the game looks too similar to the clouds in the background. It can be hard to identify them as harmful objects. The third phase has a major issue; the indicators that displays the tricks that the player has to perform does not show where the player needs to start pressing when performing said tricks. The option to switch between touch and tilt controls in the start menu has graphics that can make it hard to determine what mode has been selected.
2	Game manual or tutorial is not necessary in order to understand how to play the game.
3	The only change in the user interface occurs when the third phase starts. A new set of icons on the right side of the game now displays the tricks that the player is expected to perform.
4	The graphical style is pretty much the same. On a side-note, the start menu feels somewhat unfinished.
5	The first and second phase are easy to understand, the third phase is not immediately clear on this.
6	There is no in-game information on how to play the game during gameplay.
7	Visuals works well, sound has not yet been implemented.
8	There are no graphical elements that disturbs gameplay. Fingers can get in the way when using touch as game controller.

B.1.3 Recommendations

Gameplay

In order to increase customization it should be possible to buy power-ups such as increased life. Other visual customization options such as different costumes for the player character could also be made available. It is likely that purchase options is intended to be made available in later iterations of the game, as the player can gather resources in the game.

Difficulty increase is noticeable in the first phase, but not in the second and third phase. This is most likely a design decision. The third phase is unclear on what the player should do. A tutorial will help to alleviate this.

At the moment, there is little reason to continue playing the game. As resources can be gathered, it is likely that a shop will be made available in later stages of development. This can serve as encouragement to continue playing the game.

Mechanics

The mechanics of the game function well. The control scheme is pretty basic and easily identifiable.

Usability

Harmful clouds should have graphics that differentiate them from the clouds in the background. Animating them with lightning strikes and changing the color of the clouds in the background to white could solve this problem.

The graphics that indicate which trick the player has to perform should clearly mark where the player has to push in order to start the trick. At the moment, this graphics gives the indication to push on the opposite side of where it is needed. The menu option for switching between touch and tilt controls needs improved graphics in order to more clearly display the selected mode.

What to do in the third phase is not immediately clear. There should be a tutorial pop-up that informs the player how to play this section of the game. A redesign of the tricks mechanic may be warranted.

Depending on the size of the players size, fingers can get in the way of game-play elements when using touch controllers. An option to make the player character appear higher up on the screen could be implemented in order to alleviate this problem.

As sound has currently not been implemented fully into the game, recommendations cannot be made at the current time.

General notes

The start menu feels like part of the game, but it feels unfinished. Background graphics should be added to the game in order to make this part of the game look more appealing.

The player character had some behavior when using touch controllers that

may not be intended. When clicking on the screen to the left or right of the player character the character jumps to that position. This may be an intended function, but it also caused some unwanted movement during gameplay.

B.2 Playtesting

Test participants were observed whilst playing the game. All participants were asked to talk out loud during the play sessions. Issues they encounter during game play were written down. As it was not possible to complete the game, test participants were asked to play for 20 minutes each or until they did not want to play anymore.

B.2.1 Test Participant 1

Played for 16 minutes.

Hits the first object encountered, the weather balloon. It was not immediately clear what the second phase required the player to do, but understood this pretty quickly. The third phase was confusing and the player did not understand what to do at all. It took a while before the player saw the symbols representing what to do in order to perform the tricks.

The warning signal for the falling rock was understood as something to be avoided. However, when two came at the same time, the player did not understand how to avoid taking damage.

Complained about the control scheme as the finger covered up the player character. This caused a problem with seeing where the player character was in relation to other objects in the game. Wondered if it would be better if the player character was moved further up on the screen. There was also complaints about the character jumping to a position where the player pressed on the screen. Would have been OK if this always was the case, but the player complained about it being a delay before this could be done again.

After some time the player went to the main menu and found the settings for the controller scheme. Complained that it wasn't obvious where these settings were. Switch was made to using tilt controllers. Noted that it was hard to see from the graphics what type of controller scheme was selected. Found that the tilt controller was easier to use in the second phase, whereas touch was better for the first phase.

The end screen gave feedback which was not obvious to the player. Did not understand what XP and money was for.

B.2.2 Test Participant 2

Played for 8 minutes.

Starts the game without checking any start menu items. Hits the first object encountered, the weather balloon. Thought that they were meant to be picked up. Does not understand what the balloons are carrying. Starts looking for ways to get more life. Sees the life and meter interface elements. Recognizes that the

warning signal symbolizes something that needs to be avoided. Did not recognize that dark clouds are dangerous, looked too much like the clouds in the background. Initially thought that the clouds in the background were dangerous.

Liked the tilt control better than touch control. Would have quit earlier with touch than with tilt. Found it hard to see the player character when using touch controls. Was initially unsure of how to start the game. Initially thought that she had to drag the player character in order to start the game. The G+-button in the main menu was confusing and obfuscated the menu button for the controller settings.

Unsure of what to do in the third phase. Wonders how one can see the records for reaching the highest amount of meters.

B.2.3 Test Participant 3

Played for 8 minutes.

Starts the game right away without checking anything in the main menu. Waits for something to happen. Hits the first object, a descending weather balloon. Thought it was meant to be picked up. Understood that the warning indicated that something dangerous was about to happen and moved out of the way. Sees the indicator that measures the altitude reached. Finds and uses the pause button. Does not notice the health bar while playing. The test participant appears to have less control when using tilt compared to touch controller.

At one point the test participants just lets the player character float upwards without moving. This works well for quite some time and the height reached is comparable to other attempts.

Experiences that the player character jumps to another position on the screen when accidentally touching the screen in another area. This causes the test participant to start clicking on the spot where he wants the player character to move instead of dragging the player character with the finger.

Does not understand what to do in the third phase. Tries to draw patterns. The end screen is unclear, but understands the concept behind the experience points.

B.2.4 Test Participant 4

Played for 6 minutes.

Starts by testing what all the elements in the start menu does. Finds the controller settings, but is unsure of what controller scheme is selected. Understands what do to in the first and second phase. The third phase causes confusion.

Is not sure how to start the game. When the game starts, the participant thinks that the screen needs to be pushed in order for the player character to move upwards.

When switching to tilt controller, the participant notes that this feels better than touch controller.

B.2.5 Test Participant 5

Played for 11 minutes.

Starts the game without looking at any menu items. Says out loud that there is no need for help. When the game starts it takes a few seconds before the participant realizes that the screen needs to be touched in order for the player character to start ascending. Does not try to avoid the first object, a weather balloon. First and second phase is easily understood after this. The third phase is incomprehensible to the test participant. Asks if the speed increases in the second phase.

When playing, there is complaint about the finger covering the the player character when moving. Says that the hold on the phone must be adjusted in order to compensate for this. When switching to tilt controller, the participant tries to perform a finger gesture in order to start the game. Finds the tilt controller more enjoyable than the touch controller as there is no finger covering the player character.

How to perform tricks in the third phase is confusing. Intentionally dies repeatedly in order to get to the third phase to figure out how what to do there. Goes to the main menu and goes through all menu items in order to try to find any information on how to perform tricks in the third phase.

B.2.6 Recommendations

Start Menu

The button for changing controllers was not apparent to all of the test participants. Making this button larger and placing it in the middle of the screen as one of the main menu items should solve this problem.

Participants also had trouble identifying which controller setup was selected. The graphics were unclear as to which controller setup was currently selected. This can be solved by more use of color in order to enhance the affordance of the player. Having a green background for the selected controller scheme should resolve this issue.

In-game

Participants were unsure of how to start the game after pressing "Play" in the main menu. Having text which says "Touch screen to start", or something similar, should solve this issue.

Most of the participants hit the first object encountered, the weather balloon. This is a minor issue as it was understood that hitting the weather balloon was harmful. This issue is likely to be solved by indicating the weather balloon as a dangerous object when the tutorial is implemented.

The first and second phase of the game was fairly quickly understood. The third phase caused major issues as none of the participants had any idea of what to do, even after several attempts to figure out this phase. A major redesign of this section of the game is required in order to make it playable.

There were instances during gameplay where two meteors appeared at once. This caused frustration as the participants that experienced this was unsure of how to avoid taking damage. Meteors should be limited to one at a time.

One participant decided to just let the player character float upwards without interacting. The player character reached a height that was comparable to other attempts whilst doing this. This indicates that there might be a need for objects that more directly targets the player character in order to force more movement.

Controller Scheme

The participants complained about the player character being covered by their finger when using the touch controller scheme. This caused an inability to accurately see where the player character was in relation to objects in the game. If touch controller is to work well, then the player character must be placed higher up on the screen so that the player can control him by touching the area under him. This could be indicated by having a graphical object under the player character that indicates where to put the finger. This was a major issue and should be prioritized.

Another issue related to the touch controller is that the player character can jump to an area of the screen that is touched. This caused some frustration and in one case it was believed that this was an intended way of controlling the player character. Removing this functionality should be considered.

Other notes

It was somewhat unclear what the weather balloons are. It was thought by some to be a radio carried by a balloon. This was a minor issue and will most likely be solved by the introduction of a tutorial that explains what it is when encountered.

The clouds in the background are too similar to the ones that the player must avoid. The clouds in the background should be changed to white and the dangerous ones should be darker in color. They should also flash with lightning in order to enhance the appearance of danger.

Generally the tilt controller was preferred to the touch controller. This preference may have been influenced by the issue with the finger covering the player character when using touch controller. One participant indicated that it was preferable to use touch controller in the first phase and tilt controller in the second phase.

There is some indication that the player character needs to give more indication that damage has been taken. The health bar was also not noticed by all participants. A tutorial should take this into consideration.

The end screen was confusing. The information displayed here did not make much sense to the participants. No recommendation is made seeing as this refers to functions not yet implemented into the game

B.3 Post-interview

A set of predefined questions and questions based on observations during the play sessions were used in order to uncover potential issues and to illuminate issues that occurred during the play session.

The interview questions are designed to be open-ended questions ([Lazar et al. 2010](#), p. 111) when applicable. This forces the interview subject to consider the answer before giving it. Unlike the questions used in Chapter 4.1 these questions are not grouped into categories.

At the time of the user testing, a substantial redesign of the trick phase was being worked on. After the first round of questions the current state of this new design was presented to the test participant. They were then questioned about this new design.

B.4 Post-interview

Answers are reproduced in a succinct manner, not verbatim.

B.4.1 Questions and answers

1. How did you experience the game?
 1. Found it both exciting and frustrating. Liked the different phases, but did not understand the third phase.
 2. It was fine. There are many games that are pretty much the same. Interesting with a man in his bathing pants and the different objects that appeared while playing. The meter counter can be nice for people who like to compete.
 3. Hardly play games. Understood what to do in the different phases, except the third. Found it fun that things changed as the game progressed.
 4. Was a bit hard to understand the concept and the controller scheme. Visually it was fine.
 5. Very frustrating, but only because of the third phase. The rest was OK.
2. What did you like about the game?
 1. Answered in question 1.
 2. Answered in question 1.
 3. Liked the progression, specifically that you could see improvements when playing several times.
 4. Likes trying to beat the previous record (altitude in meters). Did not know how to see what the previous record was.
 5. The graphics were very nice. Did not find it very innovative, common game mechanics.
3. What frustrated you?

1. Answered in question 1.
 2. Did not understand what to do in order to start the game, there was no instruction on what to do. The descending "radios" was unclear, thought that they could be collected. The dark clouds were confusing. The third phase was completely incomprehensible.
 3. Missing the coins in phase 2 and that I could not find a way to get better at it.
 4. The controllers and that it was hard to understand what to do in the beginning. Kept holding the finger on the screen when using tilt controller, did not understand that it was not needed. Had some problems controlling the player character during the second phase in order to collect coins.
 5. Couldn't figure what to do in the third phase. Had to die in order to try to play the third phase again.
4. How did you find the variation in gameplay?
1. It was fun in the beginning. It's good with several phases.
 2. Liked that variation of dangerous elements.
 3. Did not find much variation in the game.
 4. Not much variation.
 5. Up, down and hell (the three phases). First two parts were pretty entertaining, but it was basically the same mechanics just in opposite directions. No meteors in the going down part.
5. Did you understand what you were supposed to do in all stages of the game?
1. Answered in question 1.
 2. Figured out that objects needed to be avoided. Found it a bit confusing to have three different phases in one game. Haven't played anything that had three different phases before, so it was a bit unfamiliar.
 3. Not immediately, understood what to do after a couple of rounds. Did not understand phase 3, tried to unsuccessfully find a pattern.
 4. No. Would have liked to instructions that tells the player what to do the first time playing the game.
 5. No.
6. How do you feel about the information that shows your status in the game?
1. Did not find the display showing the amount of coins gathered very useful. The displays showing height in meters and health worked well. Would like more feedback when taking damage, via graphics and/or sound.
 2. Felt that the altitude in meters and the health bar worked well until the third phase. The dots that appeared on the screen took all atten-

- tion.
3. Was not bad, but could not find how much health was left. This made it hard to understand what was needed to be killed.
 4. Yes, the feedback gave a good indication on what caused the player character to take damage and that losing all the health caused the player character to start descending.
 5. The first two stages had enough information. Not too much, not too little, just enough. It was absent in the third phase.
7. What kind of instructions/tutorial would you like to see in the game?
1. Would like to have a "Practice Mode" to test out how to play the game. Would like to see instructions appear on the screen when new elements are introduced.
 2. Would like to receive information on what to do the first time the game is played. This would include how to start, what to avoid and what to do in all phases of the game. A quick pop-up saying what is dangerous and what is not dangerous would be preferable.
 3. Would like to see text that tells the player to touch the screen in order to start the game. Not sure what else. Feels that simple games such as this should not require instructions in order to play. Needed only a couple of games in order to understand what to do, which is fine.
 4. When first playing the game there should be instructions, preferably a pop-up that visualize what to do.
 5. Do not like tutorial whatsoever. Like to jump right into the game. Would like tutorial for the third phase.
8. How did you find the progression in difficulty?
1. Felt that it got harder as the game progressed.
 2. It was fine. Did not feel that things happened too fast.
 3. Was there an increase in difficulty?
 4. Did not feel any increase in difficulty.
 5. First two phases were pretty much the same. Unsure if third phase is difficult as the controllers was the challenge there. Noticed that there were new elements that showed up as the player character reached higher altitude, but could not say if the difficulty increased as well. If there was an increase, it was very mild.
9. How did you find the controls? (Tilt vs. Touch)
1. The touch worked OK, but the player character was poorly placed. Did not like that the character jumped to a position where the screen was touched. Tilt worked better in phase 2.
 2. Tilt worked very well. The player character's movement speed was good, he moved as I expected him to move. Touch controller was

cumbersome. The fact that you had to keep the finger on top of the character made it hard to see where he was in relation to objects in the game. Felt that the touch controller lowered the control over the player character.

3. Found tilt controllers better. The player character was hidden by the finger when using touch controllers, which made it harder to see what was going on.
 4. Tilt worked well. Felt that there was more control using touch controller.
 5. The tilt controls were fine. The touch controller was limited by the finger covering the player character. Would only play the game with tilt controller.
10. How did you find the graphics/graphic style?
1. They were as expected. Was tricked a bit by the graphics in the beginning. Was a little unclear on what objects were dangerous.
 2. Found the graphics to be cool. There was nothing that looked out of place. All the elements worked well together. Consistent graphics style throughout the game, including the game menu. The only confusing part was the point system at the end of the game, but did not pay it much attention.
 3. For this kind of game the graphics were OK. Objects were well defined and easy to understand.
 4. It was fine. No irritating objects that were unclear.
 5. It was very nice. Liked the colors and theme.
11. Were you unclear on the function of any of the graphical objects in the game?
1. Answered in question 10.
 2. The descending "radios". Thought they were meant to be picked up. Calls them "radios", as there is no clear understanding of what they really are.
 3. Did not notice anything. The only thing that was unclear was how many times the player character could get hit by objects.
 4. The dots in the third phase.
 5. In the beginning it was unclear what was dangerous and needed to be avoided. Was unclear if I needed to avoid the "radios" or whatever it was (talking about the weather balloons).
12. What are your thoughts on the start menu?
1. Fit in with the style of the game. Would like to see a larger "Options"-menu.
 2. It was OK. A bit boring with only the logo. The different elements

work well. Would like to see the background resemble something from the game. Would most likely have understood how to change the controllers if that was the intention. The icon for controller settings works well.

3. Did not understand what the currencies were. Did not notice the button for switching controllers.
 4. Would have liked to see a button for instructions. A bit strange to see the collected currencies already on the start menu, more natural to have it display when the game starts. The graphical style works well with the rest of the game.
 5. Fits in good with the rest of the game. The low polygon square design is the same. Indifferent to G+, but is not a user.
13. What sort of items and/or upgrades would you like to be able to buy?
1. Would like to see the option to buy something that exchanges the hot air balloon with something else, like a rocket pack. Short time boot-items would also be nice, like a shield that protects the player character.
 2. Would like to buy something that can be used in the game. Don't usually buy things in games. Would maybe like to see the option to change how the player character looks. Maybe buy other stages with different graphics and elements.
 3. No, not that I can think of.
 4. Different looks for the player character. The ability to increase the health.
 5. Extra life. Laser to pierce through objects in order to clear the way up, but I would buy lasers for everything. Different skins for the character. Like a Santa Claus outfit.
14. For how long do you see yourself playing this game?
1. Depends on the frustration level.
 2. Maybe a few minutes a day. Probably not every day, might play a bit if bored.
 3. Not longer than what I have played now, which is quite long for me.
 4. Ten minutes.
 5. Don't really play mobile games.
15. Would you play the game again? Why/Why not?
1. Do not feel like there is much to gain from the game at the moment. Need to have more options in order to get further in the game.
 2. Yes, if the third phase was understandable. Splitting the game into three different phases makes the game more interesting to play.
 3. No, it's not something I find any joy in doing.

4. No, not really. It was not as catchy as I would like. Might have played it again if there were different stages.
5. Yes, just to try to finish a level (the third phase).

Questions regarding redesign of trick phase

1. How do you feel about this new design compared to the one you just played?
 1. It sounds significantly better.
 2. It would be easier to understand.
 3. Probably, what's in the game now is not good. I do not play much, so it's hard to tell.
 4. Yes, but unsure of what purpose it serves.
 5. Don't mind drawing patterns, just tell me where to start when drawing. New design might be better.
2. One or more gestures per trick?
 1. Requiring combos for each trick sounds more exciting.
 2. Several gestures per trick sound more interesting. One gesture per trick would be boring. Multiplier that counts up when correctly performing tricks in succession sound good.
 3. Would like one swipe per trick.
 4. More gestures per trick sounds more interesting.
 5. Multiple gestures per trick sounds better. Single gesture would be boring.

B.4.2 Recommendations

Start Menu

Generally the style of the start menu was deemed as fitting well in with the graphical style of the game. However, it was also considered to be a bit boring and feel unfinished. A background that ties more into the game was requested.

The button for changing controllers was not immediately obvious to all participants. If this button was made part of the main menu it could help to prevent this. As it stands now, it was hidden away to the side of the screen for some of the users. It was also not apparent which type of controller was selected. The selected style of controller should make use of colors in order to alleviate this. The selected style of controller should be marked green.

In-game

The first and second phase of the game was fairly quickly understood. Not all participants understood how to start the game. A tutorial would help to alleviate any minor issues encountered. None of the participants understood what to do in the third phase of the game. This phase requires a major redesign in order to be functional within the game. Most participants would like to see some sort of

tutorial that explains the elements in the game and what to do in the different phases.

The weather balloons caused some confusion as it was unclear if they should be avoided or not. A tutorial that tells the player what is dangerous should solve this problem.

The participants were split when it comes to the subject of difficulty increase in the game. Two felt that there was an increase, two didn't notice any changes and one noticed that there were new elements that showed up as there was progression in the game, but was unsure if the difficulty actually increased.

Controller scheme

Participants complained about having to have the finger covering the player character when using touch controller. This can be solved by moving the player character higher up on the screen. Most participants preferred the tilt controller, but this may have been because of the aforementioned issue. One participant did mention that it felt like touch controller had more control over the player character.

Other notes

When asked about what the participants would like to see as purchasable items in the game, the most common items were more life and visual changes to the game. New costumes for the main character were mentioned several times. One suggestion combined visual changes with mechanical, such as exchanging the hot air balloon with a rocket pack that goes faster. Other suggestions were short-time or limited use items such as a shield that prevents damage or a weapon that can shoot down dangerous objects.

Generally the participants would not play the game for much longer than what was done during the test. Most would not play the game again. The reasons for this are multiple. Firstly, there seems to be an absence of any true progression in the game. This might be solved when the in-game shop for upgrades becomes available and the experience system has benefits tied to it. Secondly, the third phase is the source of much frustration. The redesign that is in the works might help with this. It is recommended that a new round of user testing is executed once these changes have been implemented into the game in order to see what effect they have had on the user experience.

One participant mentioned that he would like to be able to see previous records in the game in order to try to beat them. This could also be expanded to be used in Google+ in order to compete with friends and other people's records.

Third phase redesign

All participants found that the new design sounds like a better approach. The majority of the participants preferred the multiple gestures per trick approach.

C Fancy Diving - Trick phase redesign

C.1 Goals

- Simplify the phase.
- Increase the pace.

C.2 QTE(Quick Time Event) mode

- QTE's with gestures and/or pressing.
- Only swipes?, random direction each time
 - up, down left, right and diagonals
 - character performs random trick for each gesture aka. gestures aren't tied to any specific tricks
 - The gestures can be executed anywhere on the screen.
- Each trick/gesture will have a timer on it, so you have to complete the trick within a short amount of time.
 - The timer will be represented by a bar displayed beneath the gesture. The bar will deplete in accordance with the timer.
 - When the bar is depleted, the trick gesture will disappear and count as 'failed'. A new trick gesture will then appear.
 - If the timer runs out while the correct trick's animation is playing, the trick will count as 'executed correctly'. However, the gesture itself MUST be executed before the timer runs out.
 - We can test having a slight delay from when the trick gesture appears to the timer starts, in case it gets too hard.
- Each trick completion yields a set amount of fancy paper as well as adding 1 to the FP multiplier.
 - If you fail a trick, the multiplier goes down either 1 or 0.5(never below 0, obviously).
- Additional notes:
 - Failing to stick the landing will no longer cause a loss of the multiplier.
 - I believe the loss of multiplier upon failing a trick will be challenging enough initially. We'll figure out through testing whether this is the case or not.

D CityCop - Facesketch - Test Results

D.1 Heuristic Evaluation

The same heuristics used in Appendix B were also used in this evaluation

D.1.1 Heuristics

Gameplay

Table 20: Heuristic rules for Gameplay

	Heuristic and Description
1	There is enough variation in gameplay to keep the player interested.
2	The gameplay elements are consistent in shape and function.
3	The difficulty of the game is well paced.
4	It is clear to the player what should be done during the different stages of the game.
5	Customization of the character through visual changes or power-ups should be available in order to help immerse the player.
6	Challenges in the game are enjoyable and keeps the player wanting to play more rather than becoming frustrated.

Mechanics

Table 21: Heuristic rules for Mechanics

	Heuristic and Description
1	Controllers are consistently mapped.
2	The player's status in the game is always identifiable.
3	The game controllers follows industry standards and conventions, with the exception of the third stage.

Usability

Table 22: Heuristic rules for Usability

	Heuristic and Description
1	The art in the game should give a clear indication of what they are and their purpose in the game.
2	The player does not require a manual or an in-game tutorial in order to understand how to play the game.
3	The user interface is consistent.
4	The main menu of the game feels like a part of the game.
5	There is enough information available in order to start playing the game as soon as it has loaded.
6	Information on how to play the game is available during gameplay.
7	Sounds and visual effects gives gameplay feedback to the player.
8	Interface elements do not interfere with gameplay.

D.1.2 Evaluation

Gameplay

Table 23: Heuristic evaluation of Gameplay

	Evaluation result
1	There is little variation in the current state of the game. Only one face is available to recreate. More faces must be added before a reliable evaluation of this criterion can be done.
2	Functionality in general does not change, but under certain selections of features there are more options. These do behave in a consistent manner.
3	There is no change in the difficulty level of the game.
4	Generally it is easy to understand the functionality of the game. It is unclear what to do in order to get a high percentage score. The likeness of the recreated face is somewhat hard to interpret.
5	There are no customization options beyond the core gameplay, which is all about customization.
6	There are few challenges and it is frustrating trying to recreate the face as it is unclear how this can be done in order to get a high score.

Mechanics

Table 24: Heuristic evaluation of Mechanics

	Evaluation result
1	Controllers works fine.
2	Apart from the end score, there is no way to know the status of how well the recreation is going.
3	Controllers follow industry standards for the most part. Some of the feature specific controllers can be misleading.

Usability

Table 25: Heuristic evaluation of Usability

	Evaluation result
1	The art works fine. All features that can be used are clear.
2	The game is generally easy to use. Some elements, like the movement functions for some of the features, are not always clear.
3	The interface is consistent.
4	There is no menu to speak of, only a start screen.
5	There is no information that explains what to do in order to recreate the face well.
6	No information on how to play the game is presented.
7	There are no sounds. Visual feedback is given when features are changed.
8	No interface elements interfered with gameplay.

D.1.3 Recommendations

Gameplay

More variation needs to be added. Firstly more faces is need in the game. The criterion for variation needs to be tested again once this has been done. At the moment this is hard to do reliably.

The lack of more faces also impact the difficulty testing. It is currently not possible to test if the difficulty rating in the game is well paced. Or if there is a difficulty change at all.

Getting a high score by just looking at the face seems unreasonably hard. There is no information on how to most accurately recreate the face and there is no mode where you can track you percentage score as you are going. Having a mode where the percentage is shown all the time could help the player to

understand what to look for and what features match with the face's features.

Mechanics

Some of the controls must be explained. Specifically the ones that are feature-specific. Beyond this, the game mechanics works consistently and understandably.

Usability

There are no instructions on how to play the game. The game is generally easy to understand, but some things should be explained. Feature-specific movement controls and how to get a good score should be mentioned on the start screen or during the first time the game is played.

D.2 Playtesting

D.2.1 Test Participant 1

Starts by trying to get the hair right. After this the participant checks out the different types of face shapes. After this a comment is made that it is quite difficult to remember the picture on the start page. Experiences problems with choosing the correct nose and ears. Notices that some of the eyes do not match at all. Does not remember eye color at all, picks brown at random. Focused on the face and did not remember clothes. Notices the "See the face again"-button right before finishing. Thought it was a bit strange as it was believed that the game was a memory game.

D.2.2 Test Participant 2

Notes that it looks easy to just start playing the game. Not quite sure what the game is all about. Finds no information on this. When starting sketching he makes a note that he did not read what the start screen said and that he probably should have done this. Tries different functions at random in order to figure out what they all do. Finishes the first sketch and then realizes what he is supposed to do (try to sketch the face on the front page).

Starts a second attempt. Notes that he probably started to quickly on the first attempt without reading the text. Finds it fun to try to recreate a face from memory. Likes this style of puzzle games. Finds it hard to remember features. Thinks that this may be good to train memory if there are many different selections of faces.

D.2.3 Test Participant 3

Checks the different skin colors. Makes a comment that he should probably have studied the face better before starting. Spends some time looking through the different customization options. Finds the "See the face again"-button after a little while.

Starts a second play through of the game. Tries to go more systematically

through the different features available in order to match them up better. Finds it hard to see what eye color the person has.

D.2.4 Test Participant 4

Comments that the start page is very clean. Can't really see anything that indicates that any action needs to be taken. Takes a while to understand what this is. The start button does not have much focus. Not sure how much time should be used to memorize the face. "Is this someone who has robbed me and I need to recreate his image?".

Once the game has started, the first thing that is looked at is the hair styles. Notes that it is quite clear what the goal is. The icons for the head shapes are too similar, hard to see a difference before clicking on them. Many details for the eyes. Wonders why the eyes can't be moved individually. Does not like that the "See the face again" image is placed on top. Would have liked to see both the image and the sketch at the same time, side-by-side. Questions if there is a reason for memorizing the face. Would have liked to see more variations of beards.

The background selection seems unnecessary. The options available for this does not represent anything in real life. Feels like it is a bit too cartoony and not serious enough. Reacts to the fact that zooming the picture changes that score. Can't see the relevance for this to happen. Feels that there were some limitations to how close to the image it was possible to get with the options available.

D.2.5 Test Participant 5

Starts by reading the text on the front page. Tries different eyes and face shapes. Remembers that it was a squarish face. Says that the guy in the image did not have glasses. After this the participant looks for different skin colors. Definitely white male Caucasian. Not much difference in the nose shapes, just light and perspective.

Notices the "See the face again"-button. "If I click this button I loose, right? It resets what I have done?" Does not feel that the blueprint looks anything like the picture.

Tries a second time and this time clicks the "See the face again-button. Says that it is weird that you can see the face without any repercussion when the front page clearly says "memorize".

D.2.6 Other notes

The scores from the test was in the 17-21% range. One scored 7%. The lowest score was made by the person that used the longest time to in order to recreate the image.

D.2.7 Recommendations

All participants understood how to play the game, but there was some question as to why you would play it. No context is given for what you are doing in the

game. Having a start screen that explains the reasons for this should be presented to the player. The start screen is also mostly ignored. It is probably not a good idea to have the face that needs to be remembered on the start screen. It should not be presented before the player chooses to start the game, this could help to convey the message that you need to study the face more.

It was not clear to all participants that they could view the face again after the game had started. The button for this should be changed to draw more attention. This should also be mentioned on the start page. Those that did find and use this button thought it was strange that you were asked to memorize the face when you could view it when you want to. One participant thought it would detract from his score if it was used. This seems like something that should be considered to be added to the game.

One participant mentioned that it is hard to see the eye color of the person. Making it possible to zoom in on the face might help this, but it might also remove some of the difficulty and realism of recreating a face. If this is considered, then it should be tested once it is implemented.

Background selection was only used by one participant, and it was noted by this person that it seemed out of place. Is this selection necessary? Consider removing it.

The participant that spent the most time on recreating the face also got the lowest score, by a significant amount. It is hard to say why this happened, but the selection of features that represent the face may not be the best at the moment. This should be looked at as it is the core element of the game. If this does not work in a way that is satisfying to the player, then the game is not working as intended.

D.3 Post-interview

All test participants were interviewed right after they had finished playing the game. Their answers here outlines the general impressions. It is not reproduced verbatim.

D.3.1 Questions and answers

1. How did you experience the game?
 1. It was hard, thought that it was a memory game. The game itself was fine. It was interesting seeing the score at the end. Did not initially see that I could see the face again.
 2. Would have liked to see some more information on the start screen. I did skip the start screen the first time, so there is some information there, but more is welcome. Liked the potential for a memorization style game. Fun when I realized that I had to recreate something I had just seen.
 3. Experiences it as a memory game. Would probably not have been in-

terested in trying to find and play this kind of game. Found it a bit hard to select certain features such as eyes. The graphics were a bit too "cartoony" in order to get a precise enough representation.

4. Thought it was something serious. That I had to sketch someone who had broken the law. The background selection made it feel less serious. If it is not meant to be serious, then I don't know why I played it.
5. It was an interesting game. I enjoy graphical challenges and memory games. Some of the functions I did not fully understand (talks about the functions for moving facial features around - *editor note*).

2. What did you like about the game?

1. I liked that there weren't too many options. You can explore the options in order to figure out what works. Not too many menu items. It was a bit like character customization in games. Something I don't like spending a lot of time on. It was intuitive, with the exception of the "see the face again"-button. Liked the score comparison in the end.
2. That you have to use your head.
3. I liked that you can test your memory.
4. It was easy to just start playing the game. Quickly understood how everything worked. There were some options I did not use as I was focused on other parts and did not think that they were important.
5. I liked the memory aspect. I generally like brain teasers. I found the premise interesting as soon as I looked at the title.

3. What frustrated you?

1. Nothing really frustrated me. The only thing was that I had some trouble remembering the face.
2. I don't think I encountered anything that frustrated me. The menus are easy to use.
3. The start page. I just skimmed through the next and noticed the start button and clicked it. There should be a better indication of what you are expected to do in the game. A new page should show the picture you are supposed to memorize after clicking the start button. The title says memorize, but you can see the same picture while playing.
4. That I was not able to do everything that I wanted to do. Could not adjust the eyes in the way I wanted. Wanted to be able to adjust the eyes individually. Also, the background made me question why I played it.
5. Did not understand the "See the face again"-button. No indication on what happens when it is used, do I loose points? Did not understand what the arrows under the feature selection does.

4. How did you find the variation in gameplay?

1. You do mostly the same thing all the time. Choosing different features. So, not really much variation.
 2. It was a bit short with only one face available. Would have liked to see scores for recreation of different faces in order to compare.
 3. There were a large enough selection of features. The graphics are perhaps a bit too generic.
 4. Did not think that there was much variation. Could change many features, but it did not feel like there was much of a difference.
 5. There was not much variation. It might be better with a larger selection of faces. Maybe something that forces you to make more use of the facial feature movement and rotation buttons.
5. Did you understand what you were supposed to do in all stages of the game?
1. I thought so. The different shapes made sense, as did the different functions. Was mostly focused on the selection part, not so much on how to change the placement. Did not always think about the color selection.
 2. Not during the first attempt, but I realized what to do for the second attempt. I like to just jump into games and figure out things as I go.
 3. The interface was very easy to use.
 4. Yes, I understood how to do thing, but not really why I was doing it.
 5. No, as mentioned before.
6. How do you feel about the information that shows your status in the game?
1. I only had my memory, and the sketch. That was OK when I thought it was just a memory game.
 2. Did not notice anything that shows my status in the game.
 3. The picture that you are supposed to remember should perhaps be a bit different, especially since you can select backgrounds.
 4. Had no idea. I was able to see the picture and then think that I had an idea, but this proved to be wrong as the end result was way off.
 5. Did not get enough information on how some of the buttons worked. Interpreted the "See the face again"-button as something that makes you loose points.
7. What kind of instructions/tutorial would you like to see in the game?
1. The only thing I Would have liked some information on was the "See the face again"-button. Some explanation that you could use that button in order to see the face again. The rest was self-explanatory. Some general information on what the game is about would be useful.
 2. It is a bit too easy to just skip the first page, does not appear to be part of the game itself. Because of this I did not try to remember the face.

The in-game menus and functions are easy to use. Could perhaps have an image that explains the different functions.

3. It is a very simple game. Maybe have a start page with an image that is not used in the game itself which changes into the sketch representation of how it should look.
 4. Would have liked to see more information on the start page. Like what the purpose of the game is, beyond just memorize the face, like the fact that you were rated at the end.
 5. I generally do not like tutorials. Maybe some pop-ups that explains what the different buttons do. Probably do not need it for all of the buttons.
8. How did you find the difficulty level?
1. Pretty hard. If I knew that I could see the face again it would have been much easier. Again, based on the fact that there aren't too many options. It is possible to try out the different choices.
 2. Might be a bit too easy if you can see the face again. Can be okay if you get point retracted from the score when looking at the picture. Maybe have a time limit on how long you can see the picture.
 3. I have a bad memory. It was of average difficulty. The hair seemed a bit random on the end screen. Did not look like the hair in the picture. It was also hard to see what eyebrows should be selected.
 4. Not hard to get through the game. Getting a high score seems to be very hard. Spent a lot of time and energy on creating a sketch.
 5. It was challenging. I find most things that deal with short-term memory to be hard.
9. How did you find the controls?
1. It worked well. Made sense with the interaction.
 2. Easy to understand and use.
 3. Easy controllers. Did not see that there were sub-menus under the different categories of features. Maybe make the tabs stand out a bit more.
 4. They were OK. Would have liked the option to drag facial features by using the mouse pointer instead of just the buttons under the feature selection.
 5. Menus and sub-menus were very intuitive. Feedback was immediate. Thumbnails for the different features were generally pretty good.
10. How did you find the graphics/graphic style?
1. I think it looks good. The user interface is very elegant and straightforwardly. I personally like the flat design. The characters are OK. The cartoon look works well.

2. It was good. Lots of different shapes of faces can be created. Did not notice the controllers that can move the different facial features.
 3. Not a fan of this style of graphics. Too "cartoony". Seems like old style *Yahoo!* and *Moteplass* type game graphics.
 4. Not sure. Can't put the graphics into any kind of context for the purpose of the game. Hard to make use of the graphics in order to recreate the image.
 5. It is a simple flash/mobile game style. It could be better. Did not hinder gameplay.
11. Were you unclear on the function of any of the graphical objects in the game?
1. Not while I was playing. Focused on what I was doing and did not pay much attention to the buttons I did not use. Like the controllers for zooming and moving features around. Had to test them before I understood what they did. The zoom button for scaling did not make sense. Would have tested the functions later in order to figure out what they do.
 2. Not really. The only thing that was unclear was what I was supposed to do, but that was because I did not read what the start page said. The detailed controllers for the facial features may not be very intuitive, but you see what they do when you press them.
 3. Did not notice what the color circle did until I clicked on it. It seems a bit excessive with this amount of colors.
 4. The background selection. I understood that they were backgrounds, but not the reason for them being there.
 5. The arrows and zoom functions. The RGB color selection adds too much complexity.
12. For how long do you see yourself playing this game?
1. Maybe two-three sessions. Not sure. Personally I would probably not have played it if there wasn't a reason for playing it.
 2. Might play it for 10-15 minutes a day, if I felt that this is something that may help memorization.
 3. I would probably only play through it once.
 4. As much as I played it now. I have no reason for playing it.
 5. 10-20 minutes, it depends on the increase in challenge.
13. Would you play the game again? Why/Why not?
1. Would not have played it again. I don't like customizing looks in games. It's not entertaining to make adjustments to a face.
 2. Yes, if there were more faces. I like memorization games.
 3. I do not feel the need to make any improvements to what I did in the

game.

4. No, there's nothing motivating me to play it.
5. With more faces, yes.

D.3.2 Recommendations

The aspect of it being a memory game seems clear, but not the context for actually playing the game. This was expressed several times during the interview. The start screen needs more information on what the game is about in order to add context. The face to be memorized should also not be on the start page. This should show up later in order to make it clear that it is part of the game and not just part of the start page.

The memory aspect was generally enjoyable, but being able to see the face again was confusing. There needs to be some explanation of this or there needs to be gameplay mechanics added to the "See the face again"-button. Loosing score when using it could be a solution.

A lack of variation in faces was noted as an issue. This is a limitation in this build of the game. For next round of testing more face should be added so that the variation aspect can be explored more.

The participants were split on how well they perceived the graphics. Some liked it, but others felt that it was too cartoony and generic. It was mentioned that this style of graphics made it hard to recreate the face as it bore little resemblance to the original face. Not knowing what kind of resources are available for the graphics part of the game, it is hard to make a recommendation on what to do with the graphics. There at least needs to be work put into trying to make the blueprint at the end look more like the face that is presented to the player in order to make it easier to understand how to recreate faces.

Some of the controllers for specifics facial features were unclear. They should either be made unnecessary or an explanation for their function needs to be presented to the player. The RGB color selection should be removed as it presents the player with far too many color variations.

E CityCop - Indentiface - Test Results

E.1 Heuristic Evaluation

The heuristics used are the same as those used in Appendix B and Appendix D.

E.1.1 Heuristics

Gameplay

Table 26: Heuristic rules for Gameplay

	Heuristic and Description
1	There is enough variation in gameplay to keep the player interested.
2	The gameplay elements are consistent in shape and function.
3	The difficulty of the game is well paced.
4	It is clear to the player what should be done during the different stages of the game.
5	Customization of the character through visual changes or power-ups should be available in order to help immerse the player.
6	Challenges in the game are enjoyable and keeps the player wanting to play more rather than becoming frustrated.

Mechanics

Table 27: Heuristic rules for Mechanics

	Heuristic and Description
1	Controllers are consistently mapped.
2	The player's status in the game is always identifiable.
3	The game controllers follows industry standards and conventions, with the exception of the third stage.

Usability

Table 28: Heuristic rules for Usability

	Heuristic and Description
1	The art in the game should give a clear indication of what they are and their purpose in the game.
2	The player does not require a manual or an in-game tutorial in order to understand how to play the game.
3	The user interface is consistent.
4	The main menu of the game feels like a part of the game.
5	There is enough information available in order to start playing the game as soon as it has loaded.
6	Information on how to play the game is available during gameplay.
7	Sounds and visual effects gives gameplay feedback to the player.
8	Interface elements do not interfere with gameplay.

E.1.2 Evaluation

Gameplay

Table 29: Heuristic evaluation of Gameplay

	Evaluation result
1	There is little variation in gameplay at the moment, the same thing is repeated for every play through.
2	Gameplay elements change in version 1.2 when swiping is introduced. This is not explained and there are not indicators for this happening.
3	It is hard to note if there is any change in difficulty level.
4	Version 1.0 and 1.1 are clear on what to do. Version 1.2 does not give any information on how to change the displayed face.
5	There are no options for customization, nor are there any power-ups.
6	The memory game aspect is enjoyable, but there is little motivation for continued gameplay.

Mechanics

Table 30: Heuristic evaluation of Mechanics

	Evaluation result
1	Controllers change dramatically in version 1.2 without any explanation.
2	The time limit is clearly displayed and it is understandable.
3	Controllers follows conventions used for mobile phones.

Usability

Table 31: Heuristic evaluation of Usability

	Evaluation result
1	There are no encounters of art that is not understandable.
2	No manual is needed, but there should be instructions on how the swiping works in version 1.2.
3	The user interface changes in version 1.2, apart from that it is consistent.
4	The main menu is consistent with the rest of the game.
5	It is easy to just start playing the game, with the exception of 1.2.
6	There are no in-game instructions.
7	There are no sounds in the game. Feedback on the players selection is given.
8	No interface elements interrupts the gameplay.

E.1.3 Recommendations

Gameplay

More variation in gameplay may be needed in order to keep players interested. The difficulty level needs to scale better. There is no experience of a change in difficulty level at the moment. Having short time to memorize the face and shorter time to pick the correct face are options that could be implemented. Increasing the amount of faces and making them more similar may also help to give a better feeling of an increase in difficulty level.

Sudden changes in gameplay in version 1.2 is not explained and it does not work as well as 1.0 and 1.1.

More motivation for playing is needed. There is no consistent scoring system and no rewards beyond choosing the correct face. Achievements could be added for things like picking a certain amount of correct faces in a row.

Mechanics

The controllers suddenly change in 1.2. This needs to be explained or changed.

Usability

Instructions or arrows that indicates how to swipe should be added to version 1.2 in order to make it more understandable.

E.2 Playtesting

Three different variations of the same game were tested. The different versions are labeled as 1.0, 1.1 and 1.2. Version 1.0 and 1.1 have four rounds of play while 1.2 have two rounds.

Comments the player makes are noted behind the result for the specific round. When a correct choice it made this is noted "Found", when a wrong choice is made this is noted "Not found".

E.2.1 Test Participant 1

1.0

1. Not found - Realized how the disguise works after selecting the wrong character.
2. Found - Started looking for specific features. Looks at the shape of the face.
3. Found - Looked for someone with a matching chin, narrow.
4. Found - Says out loud the features that are distinct and easy to remember.

1.1

1. Found - Notes that it is easy to spot when the first one is the correct choice.
2. Found - Only one have a matching nose.
3. Found - Looked for someone with a matching chin. Makes a note that there is no timer. Thought it may have disappeared when scrolling. (No timer in this version of the game).
4. Found - Looked for a flat nose.

1.2

1. Found - Did not realize how to swipe. Would not have realized how it worked without being told. Looked for a flat nose. Recognized that there is a loop with the same pictures.
2. Found - Makes a comment that he would like to be able to skip the face memorization when he feels that he knows enough. Would have liked a "continue"-button. Says the it is usually enough to remember the shape of the chin and nose.

E.2.2 Test Participant 2

1.0

1. Found - no comment
2. Not found - no comment
3. Found - no comment
4. Not found - Timer expired, did not realize that it happened.

1.1

1. Found - Tries to click on the picture to be memorized in order to try to skip it.
2. Not found - no comment
3. Found - no comment
4. Found - no comment

1.2

1. Found - Did not realize how to change the picture until told that swiping was needed to get to the next picture. Said that this needs to be shown.
2. Found - Makes an comment that the selection of faces are becoming recognizable.

E.2.3 Test Participant 3

1.0

1. Found - makes a comment that makes it clear that the people are wearing disguises.
2. Found - no comment
3. Found - no comment
4. Found - no comment

1.1

1. Found - no comment
2. Found - no comment
3. Found - no comment
4. Found - no comment

1.2

1. Found - Tries to click on the face that is to be memorized in order to try to skip to the next part of the game.
2. Found - no comment

E.2.4 Test Participant 4

Starts by reading the instructions out loud.

1.0

1. Found - No comment
2. Found - Feels that the timer is too short. Not enough to thoroughly scroll through the faces. It is probably intended.
3. Found - Noticed that there were more faces this time.
4. Not found - This one was hard.

1.1

1. Found - Just luck that this was found. Accidentally pressed the first face.

2. Found - Asks why you cannot skip the memorization part when you feel that you know it well enough. Also notices that there is no timer in this version.
3. Found - "I cannot see the timer, is it gone?"
4. Found - "Faces seem to be in random order."

1.2

1. Not found - "Why swipe? It is very chaotic. I get no overview of the different faces."
2. Found - No comment

E.2.5 Test Participant 5

Starts by reading the instructions on the start screen.

1.0

1. Found - Thought that the character was gonna be in a crowd, not one by one. Realizes that it is just a disguised character.
2. Found - No comment
3. Not found - Says out loud the features that will be memorized. Nose, eyes, lips.
4. Not found - No comment

1.1

1. Found - No comment
2. Found - Having no timer makes me feel better.
3. Found - No comment
4. Found - Wants to be able to skip the first image.

1.2

1. Found - Ah, like Tinder.
2. Found - No comment

E.3 Recommendations

Some participants made it clear that they were looking for specific features. This is an issue that can be solved by having more faces with some of the same or similar features. This can make it harder to spot the correct face in certain instances.

It was unclear how to use the swipe-function. All participants had to be told before they were able to use it. This could be solved by having arrows or other indicators that tell the player what to do.

The participants expressed a dislike for the swipe-version, as they felt that they had less overview of the faces.

Participants expressed a need to be able to skip the face they have to memorize when they feel that they are ready to proceed.

E.4 Post-interview

All test participants were interviewed right after they had finished playing the game. Their answers here outlines the general impressions. It is not reproduced verbatim.

E.4.1 Questions and answers

1. How did you experience the game?
 1. Did not realize what to do in the first round, but realized quickly what to do. The game was fine, but too easy. Found a pattern that I could use, which removed the challenging part of the game. Only needed to remember the nose and the chin, which isn't hard. Would have been harder if there were more variables.
 2. It did not seem to be very hard. Even though I did not get everything right I just needed to look for some specific features, like the eye color and the shape of the nose. Would have liked to skip the timer when memorizing the face. I should be able to skip forward when I feel that I remember what I need to remember.
 3. It is the kind of game that I may have played once, should I happen to come across it on a web page.
 4. It is easy to get into and just start playing. Understood immediately what I am supposed to do.
 5. Interesting, I liked it. Easy mechanics. It did not take long to understand what to do. Thought that it was gonna be more of a "Find Waldo"-style game.
2. What did you like about the game?
 1. It was simple and understandable. A bit too easy perhaps.
 2. I like memorization games.
 3. I did not like it.
 4. It was easy to understand the purpose of the game. Could just start an immediately understand how to play it. Nice to scroll up and down in order to view the faces.
 5. I like visual cues-based games. It is entertaining to me.
3. What frustrated you?
 1. Nothing really frustrated me.
 2. That I could not skip the face, as I mentioned earlier. I felt like the timer took some of my attention when trying to select the correct face. The swiping in the last game was annoying.
 3. Nothing was really frustrating, it just a kind of game that I probably would not have played. Not interesting.
 4. Could not skip the memorization part when I felt that I knew it well

- enough. The swipe-version was frustrating, did not understand the order in which the different faces appeared. Did not make sense.
5. Nothing much. Misunderstood the information on the title screen a bit (see answer to question number one - *editor note*). But even if my expectations were different, I changed my perception of the game as soon as I started playing it.
4. How did you find the variation in gameplay between the three versions?
 1. Did not notice much of a difference. Better with timer, as it proved more pressure and challenge. Liked scrolling better than swiping. Had a better overview of the different faces.
 2. Did not like the swipe version. Did not notice that there was much of a change. Would prefer to not have a time when trying to find the correct face. The timer during the memorization phase should be enough.
 3. Did not notice any variation apart from the need to swipe on the last game.
 4. Liked that you had a timer compared to the ones that did not have a timer. Liked that there was an increase in difficulty level.
 5. Liked the time pressure, but it may be too short. Did not like the swipe-version, would have been even worse with time restriction.
 5. Did you understand what you were supposed to do in all stages of the game?
 1. Yes, with the exception of the first round. Understood that I had to recognize faces, but thought that I would have to pick out a face from a crowd rather than from a list. This confused me a bit.
 2. Yes, it was pretty straightforward what I had to do. The only exception was the swipe-version.
 3. Yes, with the exception of the swipe-version.
 4. Yes.
 5. Yes.
 6. How do you feel about the information that shows your status in the game?
 1. Yes, the timer gave a good indication on how much time was left.
 2. I did not notice the time limit, as I was too busy trying to find the correct face. Maybe there should be a pop-up that says "Time up!" or something similar. I did not understand why the game suddenly ended.
 3. I felt that I got the information that I needed in order to play the game.
 4. Yes. Got a score when the correct one was selected. The color choice of red for wrong and green for correct answer was good.
 5. Yes.

7. What kind of instructions/tutorial would you like to see in the game?
 1. Maybe a little more information on what will happen on the start screen. Maybe have a first round that shows you what to do.
 2. Did not feel that I needed any more information than I was given.
 3. Maybe have a screen shot that shows a picture of what to look for. How to compare the first picture to the rest.
 4. Some tips for what to look for in order to remember facial features. Started looking for eye colors, but that did not work all the time.
 5. None, really. I do not think it needs any.
8. How did you find the controls?
 1. The controllers worked as expected.
 2. They were easy to use. Would have liked a "Skip" or "Continue"-button on the picture that needed to be memorized.
 3. The controls were good. Very easy to understand. It was easy. At least with the selection of pictures used here. I only looked for some specific features and patterns. For the most part, only one person had the correct eye color. When it was the same I looked for the shape of the mouth.
 4. Controllers were good. Scrolling was better than swiping. Also, it was better to be able to press the image rather than a button below it.
 5. Very phone-like. Nothing really bad. The swiping was a bit ...
9. How did you find the graphics/graphic style?
 1. It was fine. Simple and clear.
 2. It was a simple and nice style.
 3. I do not like this type of graphic style, but it worked well for this game.
 4. The graphics were OK. When I read the instructions on the front page I thought that I were looking for a person in a crowd.
 5. It fits the game and the platform.
10. For how long do you see yourself playing this game?
 1. Not long, only a couple of rounds. As I found a strategy early on, there was not much point in continuing when the challenge disappeared. I knew that I would find the correct person.
 2. Maybe two-three minutes a week.
 3. Not more than what I have played.
 4. I could probably play it a few more times. Not much motivation for playing the game.
 5. Maybe half an hour. I rarely play mobile games. When I do it is when I am waiting for something.
11. Would you play the game again? Why/Why not?

1. No. Needs more challenge. Might have played it again if this was the case. Not my type of game. Might have played it again if there was a focus on increasing your memorization capacity. Scoring system that can show an improvement.
 2. Yes. It's a bit hard to explain why. It is fun to see how much you can remember. It is maybe a bit too easy to remember the features.
 3. No, not my kind of game.
 4. I would play it again to see if I could improve my ability to recognize the correct face.
 5. Yes, I liked it. I liked the merging of a visual memory game with a visual cue game. Memorize a face and then find it despite minor differences.
12. Did you like four or six faces the best?
1. Did not notice that there was a difference in the number of faces shown. I felt as if there was no difference in the challenge with more faces. If there had been more persons with the same face and chin I would have had a more difficult time of selecting the correct one as I would have had to remember a third feature.
 2. Did not notice that more faces appeared. The most important is that there should be more faces that shared features with the face that you are looking for. More faces may make it more difficult, but as said, they should look more alike.
 3. I think there should be more faces to choose from as the difficulty level increases.
 4. It was fine with six faces. The increase in amount of faces should be part of a difficulty increase.
 5. I think the amount should increase as you get further in the game.
13. Was the time limit beneficial or detrimental to your enjoyment?
1. I liked having the time limit.
 2. Did not like the time limit.
 3. I preferred having a time limit as it gives more pressure. Maybe have different time limits for different difficulty levels and for different amounts of faces.
 4. Time limit provides pressure, it worked better.
 5. Detrimental right now because it was too short, but you need a time limit. Maybe change depending on the amount of faces.
14. Do you prefer portrait or landscape?
1. Prefer portrait. The game scales badly.
 2. Portrait is preferred.
 3. It is easier to play in portrait.

4. Portrait, landscape did not scale well.
 5. It was horrible with landscape.
15. Did you prefer scrolling or the swipe-function?
1. Scrolling worked better. Felt I had more overview of the different faces.
 2. I felt that I had a better overview when using scrolling.
 3. Scrolling, it was easier to get a good overview of the different faces.
 4. Scrolling, swiping was tedious. With time limit it would be even worse. Lost overview of the faces.
 5. Scrolling, it was faster and you could see more faces.
16. How do you feel about a system where you can only get a new face when swiping, you cannot go back to see a previous face?
1. Might make it more challenging, maybe too challenging. Would have been interesting to test. Might force you to think of more features in order to be sure.
 2. I think that may be extremely hard. Some faces may look very similar, so it may be easy to select the wrong one. You might end up blaming the game when you pick the wrong one. In your own head it may look like the correct one. what I did in the game.
 3. Might be interesting for a higher difficulty level. Probably not a good idea to use immediately as it is probably too hard for that.
 4. Could be a variation on gameplay. Would have been strange if it was the first game, a bit too steep increase in difficulty level.
 5. Adds another flavor to the game, not bad. Might be more of a gamble.

E.5 Recommendations

The game comes across as being too easy in its current state. Specific features were being looked for. Faces that have more similar features can be used in order to increase the difficulty level of the game. This was more important than just increasing the amount of faces.

Being able to skip the face to be memorized came up several times during the interviews. This functionality should be implemented as it was found to be frustrating having to wait for the timer to run out.

Some participants initially thought that they would have to identify the face in a crowd of other faces. This was based on the text on the first screen of the game. This is a minor issue as they said that they understood what to do as soon as the game started. Rewriting the text on the first screen of the game will help to alleviate this problem.

The swipe version of the game was universally disliked. Participants did not feel like they had a good overview of the faces in the rotation. This version of

the game should not be reused in the end product as it is now. On the subject of the variation mentioned in question 16, most of the participants were positive. Making this variation of the feature possibly viable for high difficulty levels.

Some participants would like to see some information on what to look for when trying to remember a face. Having instructions or a tutorial that point this out should be added to the game.

All participant would like to see a time limit used in some way. This created more of a challenge and pressure. Some suggested that it should vary depending on the amount of faces and the difficulty level.

With the current state of the game, no participants preferred the landscape mode over the portrait mode. If landscape is to work it need to scale far better than it currently does. The functionality might not even be needed on mobile phones. Tablets could be tested for this, but some changes to how the content is presented should still be considered.